

STIC-Biotech/ChemLib

85802

From: Jiang, Dong
Sent: Friday, January 31, 2003 7:16 PM
To: STIC-Biotech/ChemLib
Subject: 09/333,159

RECEIVED

FEB -3 2003

STIC-BIOTECH/ChemLib
(STIC)

Please search SEQ ID NO: 47

Please send results on paper to Dong Jiang in 10D-08 (mail stop CM1-10D19).
Thank you very much.

Dong Jiang (78243)
703-305-1345
U.S. Patent and Trademark Office
Art Unit 1646
dong.jiang@uspto.gov
CM1-10D08
Mail stop: CM1-10D19

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 2/3/03
Date Completed: 2/3/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: 1 _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 02 _____
WWW/Internet: _____
Other (specify): _____

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, **.rnpm** and **.rnpn**

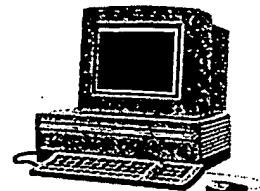
Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, **.rapm** and **.rapn**

The Pending database search results should not be left in the case because they contain data that is confidential.

BioTech-Chem Library

Search Results

Feedback Form (Optional)



Scientific & Technical Information C

The search results generated for your recent request are attached. If you have any questions or comments (compliments or complaints) about the scope or the results of the search, please contact *the BioTech-Chem searcher* who conducted the search *or contact*:

Mary Hale, Supervisor, 308-4:
CM-1 Room 1E01

Voluntary Results Feedback Form

➤ *I am an examiner in Workgroup:* (Example: 1610)

➤ *Relevant prior art found, search results used as follows:*

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ *Relevant prior art not found:*

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Search results were not useful in determining patentability or understanding the invention.

Other Comments:

Drop off completed forms at the Circulation Desk CM-1, or send to Mary Hale, CM1-1E01 or mary.hale@uspto.gov

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:48:09 ; Search time 18 Seconds
(without alignments)
691.439 Million cell updates/sec

Title: US-09-333-159-47
Perfect score: 2247
Sequence: 1 MLETTSRQWIVSHRMEMLL.....IHLMQQETNLSQGRCEAVL 4233

Scoring table: BLOSUM62

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

```

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database : Issued Patients AA:*
1 : /cgn2.6/pdatata/1/aa/5A.COMB.pep.*
2 : /cgn2.6/pdatata/1/aa/5B.COMB.pep.*
3 : /cgn2.6/pdatata/1/aa/6A.COMB.pep.*
4 : /cgn2.6/pdatata/1/aa/6B.COMB.pep.*
5 : /cgn2.6/pdatata/1/aa/PCrUs.COMB.pep.*
6 : /cgn2.6/pdatata/1/aa/PCrUk11esl.pep.*
```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1282	57.1	392	4	US-09-820-001-4	Sequence 17, Appl1
2	1161	51.7	377	2	US-08-237-108-17	Sequence 14, Appl1
3	1161	51.7	377	2	US-08-073-674-17	Sequence 17, Appl1
4	1150.5	51.2	379	1	US-08-227-108-18	Sequence 18, Appl1
5	1150.5	51.2	379	2	US-09-073-674-18	Sequence 18, Appl1
6	1141.5	50.8	379	1	US-08-227-108-3	Sequence 3, Appl1
7	1141.5	50.8	379	2	US-09-073-674-3	Sequence 3, Appl1
8	1141.5	50.8	380	1	US-08-237-108-5	Sequence 5, Appl1
9	1141.5	50.8	380	2	US-09-073-674-5	Sequence 5, Appl1
10	1127	50.2	378	4	US-09-166-489-2	Sequence 2, Appl1
11	1098.5	48.9	380	1	US-08-227-108-16	Sequence 16, Appl1
12	1098.5	48.9	380	2	US-09-073-674-16	Sequence 16, Appl1
13	1031	45.9	395	4	US-09-820-001-2	Sequence 2, Appl1
14	105	4.7	430	1	US-08-484-105-22	Sequence 22, Appl1
15	105	4.7	430	1	US-08-444-106-22	Sequence 22, Appl1
16	102	4.5	346	2	US-08-602-359A-34	Sequence 34, Appl1
17	98.5	4.4	388	1	US-08-232-519-2	Sequence 2, Appl1
18	98.5	4.4	388	1	US-08-456-556-2	Sequence 2, Appl1
19	91	4.0	298	4	US-09-355-166-6	Sequence 6, Appl1
20	90	4.0	277	4	US-09-424-349A-6	Sequence 6, Appl1
21	89.5	4.0	624	2	US-08-756-317-9	Sequence 9, Appl1
22	89	4.0	494	4	US-09-134-001C-4A75	Sequence 4A75, Appl1
23	89	4.0	1674	2	US-08-968-542C-12	Sequence 12, Appl1
24	88.5	3.9	869	2	US-08-483-101-15	Sequence 15, Appl1
25	84.5	3.8	338	2	US-08-602-359A-40	Sequence 40, Appl1
26	84.5	3.8	935	1	US-07-707-367-2	Sequence 2, Appl1
27	84	3.7	370	4	US-09-222-938A-64	Sequence 64, Appl1

28	84	3.7	719	4	US-09-3865-607-2	Sequence 2, Appl1
29	83.5	3.7	341	4	US-09-4112-600B-4	Sequence 4, Appl1
30	88.5	3.7	485	2	US-08-446-803-1	Sequence 1, Appl1
31	89.5	3.7	485	2	US-08-861-837-1	Sequence 1, Appl1
32	82.5	3.7	485	2	US-08-600-900A-12	Sequence 12, Appl1
33	82.5	3.7	485	3	US-08-683-893A-12	Sequence 12, Appl1
34	82.5	3.7	485	3	US-08-600-655-1	Sequence 1, Appl1
35	82.5	3.7	485	4	US-09-170-670-1	Sequence 1, Appl1
36	82.5	3.7	485	4	US-09-170-670-7	Sequence 7, Appl1
37	82.5	3.7	485	4	US-09-193-066-1	Sequence 1, Appl1
38	82.5	3.7	485	4	US-09-193-066-7	Sequence 7, Appl1
39	82.5	3.7	485	4	US-09-183-412-1	Sequence 1, Appl1
40	82.5	3.7	485	4	US-09-183-412-7	Sequence 7, Appl1
41	88.5	3.7	485	4	US-09-264-097-7	Sequence 7, Appl1
42	88.5	3.7	485	4	US-09-354-191A-1	Sequence 1, Appl1
43	82.5	3.7	485	4	US-09-291-022A-19	Sequence 19, Appl1
44	82.5	3.7	485	4	US-09-290-734-1	Sequence 1, Appl1
45	82.5	3.7	485	4	US-09-290-734-7	Sequence 7, Appl1

ALIGNMENTS

```

RESULT 1
US-09-820-001-4
: Sequence 4, Application US/09820001
: Patent No. 6387680
: GENERAL INFORMATION:
: APPLICANT: MERKULOV, Gennady et al
: TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
: TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
: TITLE OF INVENTION: THEREOF
: FILE REFERENCE: CL001186
: CURRENT APPLICATION NUMBER: US/09/820.001
: CURRENT FILING DATE: 2001-03-29
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4
: LENGTH: 392
: TYPE: PRF
: ORGANISM: Human
: US-09-820-001-4

```

[illegible]

Db 249 KNLNVSREFDVYLGHNPAAGTSVQDFLHWAQLVRSKGKFAFNWGSPSQNMLHYNQKTPPEYD 308

Db 184 TVKYTKSLVKNLRFIPPTMFKIIFGDKIIFYPHNFDFQFLATOVCSRETLNVIGSNALFII 243
QY 277 GGFNTNMNMSRASVYAHTLAGTSVQNTILHWSQAVNSGELRAFDAGSETKNLEKCNQPT 336
Db 244 CGFDSANLNSRLDVYSHNPAGTSVQNLHWTQAVKSGNFQAFNGSPAQNVYHFNQPT 303
QY 337 PVRYRVRTMTVPTAMTGGODWLSNPEDVKMLLSEVTNLIYHKNIPEWAHVDFIWLGDAP 396
Db 304 PPYVNTAMNVPJAVNSGGNDWLDQDDVLLPLKLSNLIYHKIILPYNHLDFIWAHNP 363
QY 397 HRMYNEIHLMOQEE 411
Db 364 QEVYNEIISMAKDK 378

RESULT 6

US-08-227-108-3

; Sequence 3, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-227-108-3

Query Match 50.8%; Score 1141.5; DB 1; Length 379;
Best Local Similarity 56.5%; Pred. No. 3.4e-112;
Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;

QY 37 PTKAVDPEAFMNISEIIHQGYPCEEYEVATEDGYILSVNRIPRGLVQPKKTSRPPVLL 96
Db 7 PT---NPEVTMNIQSMTIYWGYPAAEYEVVTEGDYILGIDRIPYGRKNSENIGRRPVAF 63
QY 97 QHGLVGASWISNLPNNSLGLFADAGFDVWNGSGNWSRKHHTLSIDQDEFWAFSY 156
Db 64 QHGLASATWISNLPNNSLAFILADAGYDVLGNSRGNTWARRNLVYSPDSVEFWAFSF 123
QY 157 DEMAREDLPAVINFILOKTOGEKIYYGVYSGQTTMGFIATSTPELAQKIKMYFALAPIA 216
Db 124 DEMAKYDLPATIOFILKKTQDKLHYVGHSGQTTIGFIATSTPNKAKRIKTFYALAPVA 183

QY 217 TVKHAKSPGTFKLLPDMIMKGLFGKKEFLYQTRFLRQ-LVYILCGQVILDQICSNIMLL 275
Db 184 TVAYTETLLKNMLVPSFLFKLIFGNKIF-YPHFFDOFLATEVCSRETVDLCSNALFI 242
QY 276 LGGFNTNMNMSRASVYAHTLAGTSVQNTILHWSQAVNSGELRAFDAGSETKNLEKCNQ 335
Db 243 ICGFDTMNLNMSRLDVYSHNPAGTSVQNLHWSQAVKSGFQAFDNGSPVQNMHYHQS 302
QY 336 TPVRYRVRTMTVPTAMTGGODWLSNPEDVKMLLSEVTNLIYHKNIPEWAHVDFIWLGLDA 395
Db 303 MPPYVNTLDMHVPJAVNWNGNDLADPHVDVLLSKLPNLIYHKKIPPPYHNLDFIWAHDA 362
QY 396 HRMYNEIHLMOQEE 407
Db 363 QPQVYNEIVSM 374

RESULT 7

US-09-073-674-3

; Sequence 3, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-073-674-3

Query Match 50.8%; Score 1141.5; DB 2; Length 379;
Best Local Similarity 56.5%; Pred. No. 3.4e-112;
Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;

QY 37 PTKAVDPEAFMNISEIIHQGYPCEEYEVATEDGYILSVNRIPRGLVQPKKTSRPPVLL 96
Db 7 PT---NPEVTMNIQSMTIYWGYPAAEYEVVTEGDYILGIDRIPYGRKNSENIGRRPVAF 63
QY 97 QHGLVGASWISNLPNNSLGLFADAGFDVWNGSGNWSRKHHTLSIDQDEFWAFSY 156
Db 64 QHGLASATWISNLPNNSLAFILADAGYDVLGNSRGNTWARRNLVYSPDSVEFWAFSF 123
QY 157 DEMAREDLPAVINFILOKTOGEKIYYGVYSGQTTMGFIATSTPELAQKIKMYFALAPIA 216

Db 124 DEMAKYDLPATIDFILKKTGQDKLHVHSGQTTIGFIAFSTNPKLAKRIKTFYALAPVA 183
QY 217 TVKHAKSPGTFKLLPDMITKGLFGKKEFLYQTRFLRQ-LVIVLCGOVILDOICSNIMLL 275
Db 184 TVKYTELLNKLMLVPSFLFKLIFGNKIF-YPHHFFDQFLATEVCSRETVDLLCSNALFI 242
QY 276 LGGFNTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDWGSFETKMLEKCNQ 335
Db 243 ICGFDTMNLNMSRLDYLSHNPAGTSVQNVLHWSQAVKSGKFOAFDQWSPVQNMHHYQS 302
QY 336 TPVRYVRDVTPTAMWTGGQDLSNPEDVKMLLSEVNTLIYKNIPEVAHWDFIWLGLDA 395
Db 303 MPYYNLTDMHVPVAVNGNDLLADPHDVLKLPNLIYHRKIPPPYNNHLDIWMADA 362
QY 396 PHRYNEIHLM 407
Db 363 PQAVYNEIVSM 374

RESULT 8

US-08-227-108-5
; Sequence 5, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allen A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-227-108-5

Query Match 50.8%; Score 1141.5; DB 1; Length 380;
Best Local Similarity 56.5%; Pred. No. 3.5e-112;
Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;
QY 37 PTKAVDPEAFNISEIIOHOGYPCPEEYEVATEDGYILSVNRIPRGLVQPKTGRSPVLL 96
Db 8 PT---NPEVTNISQMITYWCYPAEYEVVTDGYILGIDRIPYGRKNSENIGRRPVAFL 64
QY 97 QHGLVGASWISNLPNNLSGLILADAGFDVWNGNSRGNAWSRKHKTLSIDQDFWAFSY 156
Db 65 QHGLLASATNWSILNPNLSLAFILADAGYDVLGNSRGNTWARNLNYSFSDSVEFWAFSE 124

QY 157 DEMAREDPVAVINFILQKTQOEKIYVYVSGQTTMGFIAFSTMPELAQKIKMYFALAPIA 216
Db 125 DENAKYDLPATIDFILKKTGQDKLHVHSGQTTIGFIAFSTNPKLAKRIKTFYALAPVA 184
QY 217 TVKHAKSPGTFKLLPDMITKGLFGKKEFLYQTRFLRQ-LVIVLCGOVILDOICSNIMLL 275
Db 185 TVKYTELLNKLMLVPSFLFKLIFGNKIF-YPHHFFDQFLATEVCSRETVDLLCSNALFI 243
QY 276 LGGFNTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDWGSFETKMLEKCNQ 335
Db 244 ICGFDTMNLNMSRLDYLSHNPAGTSVQNVLHWSQAVKSGKFOAFDQWSPVQNMHHYQS 303
QY 336 TPVRYVRDVTPTAMWTGGQDLSNPEDVKMLLSEVNTLIYKNIPEVAHWDFIWLGLDA 395
Db 304 MPYYNLTDMHVPVAVNGNDLLADPHDVLKLPNLIYHRKIPPPYNNHLDIWMADA 363
QY 396 PHRYNEIHLM 407
Db 364 PQAVYNEIVSM 375

RESULT 9

US-09-073-674-5
; Sequence 5, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-073-674-5

Query Match 50.8%; Score 1141.5; DB 2; Length 380;
Best Local Similarity 56.5%; Pred. No. 3.5e-112;
Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;
QY 37 PTKAVDPEAFNISEIIOHOGYPCPEEYEVATEDGYILSVNRIPRGLVQPKTGRSPVLL 96
Db 8 PT---NPEVTNISQMITYWCYPAEYEVVTDGYILGIDRIPYGRKNSENIGRRPVAFL 64
QY 97 QHGLVGASWISNLPNNLSGLILADAGFDVWNGNSRGNAWSRKHKTLSIDQDFWAFSY 156

Db 65 QHGLASATNWSLNPNLSAFILADAGYDVWLGNSRGNTWARRNLYYSPDSVEFWAFSF 124
QY 157 DEMARFDPVAVNFILQKTQKQKIYVGYSGTTFAGTAFSTMPDLAQKIKMYFALAPTA 216
Db 125 DEMAKYDLPATIDFLIKTKTGDKLHVGHSGTTFAGTAFSTMPDLAQKIKMYFALAPYA 184
QY 217 TVKHAKSPCTKFLLLPDMWIKGLFGKKEFLYOTRELRQ-LVIYLCGOVILDOICSNIMLL 275
Db 185 TVKYTETLLNKLMLVPSFLFKLIFGNKIF-YPHHFFDQFLATEVCSRETVDLLCSNALFI 243
QY 276 LGGFNTNMMNSRASVYAHTLAGTSVONILHWSQAVNSGELRAFDMGSETKNLEKCNQP 335
Db 244 ICGFOTMLNMSRLDVLHSHNPAGTSVQNVLHWSQAVNSGKFAEDMGSPVQNMHHYHQS 303
QY 336 TPVRYVRDMTPTAMWTGGQDLNPNEDVKMLLSEVTNLIYHKNIPENAHVDFIWLGLDA 395
Db 304 MPYYNLTDMHVPVIAVWNGNDLADPHDVLKLLKPLNLIYHKNIPENAHVDFIWLGLDA 363
QY 396 PHRYMYNEIHLM 407
Db 364 POAVYNEIVSMM 375
RESULT 10
US-09-186-489-2
; Sequence 2, Application US/09186489
; Patent No. 6375947
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul L.
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
; TITLE OF INVENTION: Processes for Its Production and Use
; FILE REFERENCE: 5499/3
; CURRENT APPLICATION NUMBER: US/09/186,489
; CURRENT FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Kid (Goat)
US-09-186-489-2
Query Match 50.2%; Score 1127; DB 4; Length 378;
Best Local Similarity 55.1%; Pred. No. 1.2e-110;
Matches 205; Conservative 66; Mismatches 101; Indels 0; Gaps 0;
QY 40 AYDPEAFMISEIIOHGYPCPEYEVATEDGYILSVNRIPIRGVLQPKKTSRPPVLLQHG 99
Db 6 ANPEASMVWSQISWGPSEMHKVTADGYILQVYRPHGKNDANHLGQRPVFLQHG 65
QY 100 LVGGASNWSLNPNLSGLFILADAGYDVWLGNSRGNTWARRNLYYSPDSVEFWAFSDYM 159
Db 66 LLASATNWSLNPNLSGLFILADAGYDVWLGNSRGNTWARRNLYYSPDSVEFWAFSDYM 125
QY 160 ARPDLPVAVNFILQKTQKQKIYVGYSGTTFAGTAFSTMPDLAQKIKMYFALAPATVK 219
Db 126 AEYDLPSTIDFLIKRTGQKHLHVGHSGTTFAGTAFSTMPDLAQKIKMYFALAPATVK 185
QY 220 HAKSPCTKFLLLPDMWIKGLFGKKEFLYOTRELRQ-LVIYLCGOVILDOICSNIMLLGGF 279
Db 186 HTQSLFNKLALIPHEFLFKLIFGNKIF-YPHHFFDQFLATEVCSRETVDLLCSNALFI 245
QY 280 NTNMMNSRASVYAHTLAGTSVONILHWSQAVNSGELRAFDMGSETKNLEKCNQPPTVR 339
Db 246 DNKNFMSRLDVLHSHNPAGTSVQNVLHWSQAVNSGKFAEDMGSPVQNMHHYHQS 305
QY 340 YVRDVTPTAMWTGGQDLNPNEDVKMLLSEVTNLIYHKNIPENAHVDFIWLGLDAPHRM 399
Db 306 YNLTMVPIAVWSAGQDLNPNEDVKMLLSEVTNLIYHKNIPENAHVDFIWLGLDAPHRM 309
QY 400 YNEIHLMOQEE 411

Db 366 YNEIHLMAKOK 377
RESULT 11
US-08-227-108-16
; Sequence 16, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-227-108-16
Query Match 48.9%; Score 1098.5; DB 1; Length 380;
Best Local Similarity 54.6%; Pred. No. 1.2e-107;
Matches 202; Conservative 69; Mismatches 98; Indels 1; Gaps 1;
QY 43 PEAFMISEIIOHGYPCPEYEVATEDGYILSVNRIPIRGVLQPKKTSRPPVLLQHG 102
Db 10 PEVTMISOMITWYGPNEEYEVATEDGYILEVNRIPYKKNKSGNTGQRPVFLQHG 69
QY 103 GASNWSLNPNLSGLFILADAGYDVWLGNSRGNTWARRNLYYSPDSVEFWAFSDYM 161
Db 70 SATNWSLNPNLSGLFILADAGYDVWLGNSRGNTWARRNLYYSPDSVEFWAFSDYM 129
QY 162 FDPVAVNFILQKTQKQKIYVGYSGTTFAGTAFSTMPDLAQKIKMYFALAPATVK 221
Db 130 YDLPATIDFIVKTKGQKHLHVGHSGTTFAGTAFSTMPDLAQKIKMYFALAPATVK 189
QY 222 KSPGTKFLLLPDMWIKGLFGKKEFLYOTRELRQ-LVIYLCGOVILDOICSNIMLLGGFNT 281
Db 190 KSLINKLRFVPSLFKIFGDKIFYPHNFDDQFLATEVCSREMLNLCSNALFIICGDS 249
QY 282 NNMMNSRASVYAHTLAGTSVONILHWSQAVNSGELRAFDMGSETKNLEKCNQPPTVR 341
Db 250 KNFTSLDVLHSHNPAGTSVQNVLHWSQAVNSGKFAEDMGSPVQNMHHYHQS 309
QY 342 VRDMTPTAMWTGGQDLNPNEDVKMLLSEVTNLIYHKNIPENAHVDFIWLGLDAPHRM 401

Db 310 VTAMNPIAVNNGKDLADPODVGLLPKLPNLIYHKEIPFYNHLDFTIWDAPQEVYN 369

Qy 402 EIIHLMQOEE 411

Db 370 DIVSMISEDK 379

RESULT 12

US-09-073-674-16
; Sequence 16, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Julien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073.674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-073-674-16

Query Match 48.9%; Score 1098.5; DB 2; Length 380;

Best Local Similarity 54.6%; Pred. No. 1.2e-107; Indels 1; Gaps 1;

Matches 202; Conservative 69; Mismatches 99; Indels 1; Gaps 1;

Qy 43 PFAFNISIIHQOGYPCPEYEVATEDGYILSVNRIPRGLVQPKTGRSPVVLQHGVLG 102

Db 10 PEVTNISMOMITYGYNVEYEVATEDGYILEVNRIPYKKNKSGNGTQRPVVLQHGLLA 69

Qy 103 GASNINSLPNNLSGLIFILADAGFDVWNGSRGNANRKHKTLSIDQDEFW-AFSYDEMAR 161

Db 70 SATNINSLPNNLSGLIFILADAGFDVWNGSRGNANRKHKTLSIDQDEFW-AFSYDEMAR 129

Qy 162 FOLPAVINFILOKTOGKIYVYSGTGMGTFAPSTPELAOKIKMYFALAPIATVKHA 221

Db 130 YOLPATIDFIVKTKQKQLHYVGHSGQTGTFIAPSTWPSLAKRIKTYFALAPVATVYT 189

Qy 222 KSPGTFKLLPDMKIKLFGKKEFLYQTRFLRQLVYLCGVILDOICSNIMLLLGFGNT 281

Db 190 KSLINKLRFPVQSLFKFIFGDKIFYPHNFDDQFLATEVCSREMLNLLCSNALFIICGDS 249

Qy 282 NNNMSRASVYAAHTLAGTSVQNLHWSQVNSGELAFDNGSETKNEKCNQPTPVYR 341

Db 250 KNFTSRDLVYLHSHNPAGTSVQNMFWHTQAVKSGKFOAYDMGSPVQNRMHYDQSQPPYIN 309

Qy 342 VRDMTVPMTAGTGODWLSNPEDVKMLLSEVNTLIYHKNIPWAHVDFTWGLDAPHMYN 401

Db 310 VTAMNPIAVNNGKDLADPODVGLLPKLPNLIYHKEIPFYNHLDFTIWDAPQEVYN 369

Qy 402 EIIHLMQOEE 411

Db 370 DIVSMISEDK 379

RESULT 13

US-09-820-001-2
; Sequence 2, Application US/09820001
; Patent No. 6387680
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01186
; CURRENT APPLICATION NUMBER: US/09/820.001
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Human
US-09-820-001-2

Query Match 45.9%; Score 1031; DB 4; Length 395;

Best Local Similarity 51.0%; Pred. No. 1.8e-100;

Matches 199; Conservative 61; Mismatches 124; Indels 6; Gaps 4;

Qy 17 MWLLILVAYMPQNVNS---VHMPKAVDPEAFMNISEIIHQGIPCEYEVATEDGYIL 73

Db 2 MWLLITTTCLICGTLNAGGFLDLENE-VNPEYVMMNTSEIIYNGVPSEYEVATEDGYIL 60

Qy 74 SVNRIPRGLVQPKTGRSPVVLQHGVLGASNINSLPNNLSGLIFILADAGFDVWNGSR 133

Db 61 LVNRIPYRTHARTSTGPRPVYVMYHALFADNAYLNANGSLGFLADAGFDVWNGSR 120

Qy 134 GNWSRKHKTLSIDQDEFWAFSYDEMARFDLPAVINFILOKTOGKIYVYSGTGMG 193

Db 121 GNWSRKHKTLSIDQDEFWAFSYDEMARFDLPAVINFILOKTOGKIYVYSGTGMG 180

Qy 194 IAFSTMPELAOKIKMYFALAPIATVKHAKSPCTKFLLLPDMKIKLFGKKEFLYQTRFLR 253

Db 181 VAFSTMPELAOKIKMYFALAPIATVKHAKSPCTKFLLLPDMKIKLFGKKEFLYQTRFLR 240

Qy 254 QLVVYLCGOVILDOICSNIMLLLGCFNTNMMNSRASVYAAHTLAGTSVQNLHWSQAVN 313

Db 241 IASTKICNNKILWLCSEFMSLWAGSNKKMMNOSRMDVYMSHAPTSSVHNLHILKOLYH 300

Qy 314 SOELRAFDMGSETKNEKCNQPTPVYRVRDMTVPMTAGTGODWLSNPEDVKMLLSEV 373

Db 301 SDEFRAYDNGDADNKKHYNOSHPPYIDLTAMKVTATWAGGHVLTGTQDVARILPOIK 360

Qy 374 NL-IYHKNIPWAHY-DFIWLGLDAPHMYN 401

Db 361 SLSLSVLSPLEWEPTDFVWGLDAPQRMFS 390

RESULT 14

US-08-484-105-22

; Sequence 22, Application US/08484105

; Patent No. 5589341

; GENERAL INFORMATION:

; APPLICANT: STILLMAN, Bruce

; APPLICANT: BELL, Stephen P

; APPLICANT: KOBAYASHI, Ryuji

; APPLICANT: RINE, Jasper

; APPLICANT: FOSS, Margit

APPLICANT: STILLMAN, Bruce
APPLICANT: BELL, Stephen P
APPLICANT: KOBAYASHI, Ryuji
APPLICANT: RINE, Jasper

Search completed: February 3, 2003, 13:50:50
Job time : 19 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:44:29 ; Search time 40 Seconds
(without alignments)
1409.125 Million cell updates/sec

Title: US-09-333-159-47

Perfect score: 2247

Sequence: 1 MLETLSRQWIVSHRMEMWLL.....IHLMQEETNLSQGRCEAVL 423

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2247	100.0	423	22 AAB66065	Human TANGO 294.
2	2174	96.8	409	22 AAEL1931	Human CGI62 (or C5
3	2076	92.4	390	22 AAB66067	Human TANGO 294 ma
4	1991	88.6	398	23 AAU99164	Human lysosomal ac
5	1536.5	68.4	371	23 AAU99165	Human lysosomal ac
6	1289	57.4	399	22 AAB200783	Human shear stress
7	1289	57.4	399	22 AAB66061	Human lysosomal ac
8	1282	57.1	392	23 AAO18227	Human lysosomal ac
9	1191.5	53.0	449	22 AAU30498	Novel human secret
10	1166	51.9	398	14 AAR37302	RGL precursor. Or

11	1163	51.8	221	22 AAB66068	Human TANGO 294 ex
12	1156.5	51.5	395	6 AAP50322	Rat lingual lipase
13	1143	50.9	395	23 AAEL14744	Human triacylglyce
14	1143	50.9	399	23 AAU77496	Human lipid metabo
15	1141.5	50.8	379	15 AAR56870	Canine gastric lip
16	1141.5	50.8	379	17 AAU09382	Dog gastric lipase
17	1141.5	50.8	380	15 AAR56871	Canine gastric lip
18	1139	50.7	399	22 AAG67513	Amino acid sequenc
19	1135	50.6	397	23 AAG17308	Human lysosomal ac
20	1127	50.2	378	23 ABB76189	Kid goat ptegrastri
21	1119	49.8	398	7 AAP60724	Sequence of pregas
22	1119	49.8	398	7 AAP60658	Sequence of human
23	1119	49.8	398	17 AAU09383	Human gastric lipa
24	1119	49.8	398	22 AAB66086	Human lipid prote
25	1101	49.0	403	23 AAU77493	Human lipid metabo
26	1090	48.5	427	23 AAU98539	Human lysosomal ac
27	1031	45.9	395	23 AAO18226	Human lysosomal ac
28	977	43.5	365	23 AAEL17307	Human lysosomal ac
29	907	40.4	731	23 AAEL14746	Human triacylglyce
30	783	34.8	144	22 AAB66070	Human TANGO 294 cy
31	781	34.8	276	23 AAEL14745	Human triacylglyce
32	777	34.6	289	23 AAU77494	Human lipid metabo
33	745.5	33.2	280	23 AAEL14743	Human triacylglyce
34	724.5	32.2	656	22 ABG26839	Novel human diegno
35	691	30.8	233	22 AAB61608	Human protein HP03
36	663	29.5	394	22 ABB71702	Drosophila melanog
37	655.5	29.2	434	22 ABB62081	Drosophila melanog
38	639	28.4	456	22 ABB68957	Drosophila melanog
39	636	28.3	398	22 ABB63011	Drosophila melanog
40	627	27.9	311	23 AAU98540	Human lysosomal ac
41	588.5	26.2	838	22 ABB61967	Drosophila melanog
42	567.5	25.3	457	22 ABB63187	Drosophila melanog
43	559	24.9	399	22 ABB71442	Drosophila melanog
44	546	24.3	439	22 ABB63143	Drosophila melanog
45	535	23.8	416	22 ABB59328	Drosophila melanog

ALIGNMENTS

RESULT 1
AAB66065
ID AAB66065 standard; Protein: 423 AA.
XX AC AAB66065;
XX DT 30-MAR-2001 (first entry)
XX DE Human TANGO 294.

XX KW TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety.
XX OS Homo sapiens.

XX XX WO200077239-A2.
XX PD 21-DEC-2000.
XX PF 24-MAY-2000; 2000WO-US14858.
XX PR 14-JUN-1999; 99US-0333159.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
XX WPI; 2001-032313/04.
XX DR N-PSDB; AAF45131, AAF45132.

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT screening assays and diagnostic assays and for the treatment of
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT disease.

XX Claim 8; Fig 6; 359pp; English.

XX The present invention relates to TANGO or INTERCEPT proteins and coding
XX sequences (see AAF45121-P45136 and AAF45138-P45139 and AAF66031-B66057,
CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC sequences are useful for the treatment of neurological disorders such as
CC central nervous system (CNS) disorders, CNS-related disorders, focal
CC brain disorders, global-diffuse cerebral disorders and other
CC neurological and cerebrovascular disorders. The CNS disorders include
CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC autonomic function disorders such as hypertension and sleep disorders,
CC neuropsychiatric disorders, psychoactive substance use disorders,
CC anxiety, and bipolar affective disorder.

XX Sequence 423 AA;

Query Match 100.0%; Score 2247; DB 22; Length 423;
Best Local Similarity 100.0%; Pred. No. 8.5e-220;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLETLGROWIVSHRMENWLLILVAYMFQFNVNSVHMPTKAVDPEAFNMISEIIHQGYPC 60
DB 1 MLETLGROWIVSHRMENWLLILVAYMFQFNVNSVHMPTKAVDPEAFNMISEIIHQGYPC 60

QY 61 EYEYEVATEDGYILSVNRIPRGLVQPKKTGSRPVVLLQHLVGGASNMISLNPNSLGFIL 120
DB 61 EYEYEVATEDGYILSVNRIPRGLVQPKKTGSRPVVLLQHLVGGASNMISLNPNSLGFIL 120

QY 121 ADAGFDVWNGNSRGNAWSRKHKILSDODDFWAFSDDEMARFOLPAVINFILOKTGOEKI 180
DB 121 ADAGFDVWNGNSRGNAWSRKHKILSDODDFWAFSDDEMARFOLPAVINFILOKTGOEKI 180

QY 181 YYVGSOGTTMGFTAFSTPELAQKIMYFALAPIATVKHAKSPGTFLLLPDMWIKGLF 240
DB 181 YYVGSOGTTMGFTAFSTPELAQKIMYFALAPIATVKHAKSPGTFLLLPDMWIKGLF 240

QY 241 GKKEFLYQTRFLRQLVYLQGVILQDQICSNIMLLGGFTNNMNSRASVYAAHTLAGT 300
DB 241 GKKEFLYQTRFLRQLVYLQGVILQDQICSNIMLLGGFTNNMNSRASVYAAHTLAGT 300

QY 301 SVQNILHWSQAVNSGELRAFDMGSETKNLEKCNQPTPVRYVRDMDTPTAMWTGGQDWLS 360
DB 301 SVQNILHWSQAVNSGELRAFDMGSETKNLEKCNQPTPVRYVRDMDTPTAMWTGGQDWLS 360

QY 361 NPEDVKMLLSEVTNLIYHKNIPEWAHVDFIWLGDAPHRMVNEIHLMQOETNLSQGRCE 420
DB 361 NPEDVKMLLSEVTNLIYHKNIPEWAHVDFIWLGDAPHRMVNEIHLMQOETNLSQGRCE 420

QY 421 AVL 423
DB 421 AVL 423

RESULT 2
AAE11931 standard; Protein; 409 AA.

XX AAE11931;
AC AAE11931;
XX 18-DEC-2001 (first entry)
DT 18-DEC-2001 (first entry)
XX Human CG162 (or C59) lipase protein #2.
XX Human; apolipoprotein; lipase; lipoprotein receptor; ALLr; angina;
KW cardiovascular disease; lipid metabolism; myocardial infarction;
KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;

KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
KW neuroprotectant; cerebroprotective.

XX Homo sapiens.

XX Key Location/Qualifiers
FH Peptide 1..19
FT /label= Signal_peptide
FT Protein 20..409
FT /note= "Human mature CG162 (or C59) lipase protein"
XX WO200179446-A2.
XX 25-OCT-2001.
XX 16-APR-2001; 2001WO-US12529.
XX 14-APR-2000; 2000US-197137P.
XX 20-JUN-2000; 2000US-0598042.
XX 03-AUG-2000; 2000US-0631451.
XX 22-SEP-2000; 2000US-0667298.
XX 17-NOV-2000; 2000US-0714936.
XX (HYSE-) HYSEQ INC.
XX Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
PI Wang D;
XX WPI; 2001-611724/70.
XX N-PSDB; AAD19226.
XX Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
PT receptor polypeptides, useful for preventing diagnosing and treating
PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
XX Claim 10; Fig 3; 266pp; English.

XX The invention relates to polynucleotides encoding proteins CG122, CG179,
CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
CC involved in lipid metabolism and cardiovascular disease such as human
CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
CC and protein sequences are useful for treating or preventing disorders
CC associated with apolipoproteins, lipases and lipoprotein receptor (ALLr)
CC expression and for treating lipid metabolism, cardiovascular diseases
CC and thrombosis. Antibodies against these proteins are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of these sequences. ALLr polypeptides are also
CC useful for identifying agents (agonists and antagonists) that bind to
CC them and cells expressing ALLr proteins are useful for identifying a
CC therapeutic agent for use in treatment of a pathology related to
CC aberrant expression or physiological interactions of this polypeptide.
CC vectors comprising these DNA and protein sequences are also useful for
CC producing ALLr proteins. The nucleic acids and polypeptides of the
CC invention are also useful for the treatment of occlusive cardiovascular
CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
CC or intracardiac thrombosis and stroke. The nucleotides of the invention
CC are used in gene therapy. The present sequence is human CG162 (or C59)
CC lipase protein.

XX Sequence 409 AA;

Query Match 96.8%; Score 2174; DB 22; Length 409;
Best Local Similarity 100.0%; Pred. No. 2.2e-212;
Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 MEMWLLILVAYMFQFNVNSVHMPTKAVDPEAFNMISEIIHQGYPCPEYEVATEDGYLS 74
DB 1 MEMWLLILVAYMFQFNVNSVHMPTKAVDPEAFNMISEIIHQGYPCPEYEVATEDGYLS 60
QY 75 VNIIPRGLVQPKKTGSRPVVLLQHLVGGASNMISLNPNSLGFILADAGFDVWNGNSRG 134

|||||
Db 61 VNRIPRGLVQPKTGRPVVLLQHLVGGASNNISNLPNNSLGFIADAGFDVWGNRSG 120
QY 135 NANSRKHKTLSIDQDEFWAFSYDEMARDLPAVINFILOKTGOEKIYVYGYSQGTMGFI 194
Db 121 NANSRKHKTLSIDQDEFWAFSYDEMARDLPAVINFILOKTGOEKIYVYGYSQGTMGFI 180
QY 195 AFSTMPPELAQIKMYFALAPIATVYKHAKSPGTFFLLPDMMIKGLFGKKEFLYQTRFLRQ 254
Db 181 AFSTMPPELAQIKMYFALAPIATVYKHAKSPGTFFLLPDMMIKGLFGKKEFLYQTRFLRQ 240
QY 255 LVIVLCQVILDOICSNIMLLGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNS 314
Db 241 LVIVLCQVILDOICSNIMLLGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNS 300
QY 315 GELRAFQWSETKNLEKCNQPTPVYRVRDMTPTAMWTGGQDWSNPEDVKMLLSEVTN 374
Db 301 GELRAFQWSETKNLEKCNQPTPVYRVRDMTPTAMWTGGQDWSNPEDVKMLLSEVTN 360
QY 375 LIYHKNPIEWAHVDFINGLQDAPHMYNEIHLMQOQETNLSQGRCEAVL 423
Db 361 LIYHKNPIEWAHVDFINGLQDAPHMYNEIHLMQOQETNLSQGRCEAVL 409

RESULT 3
AAB66067
ID AAB66067 standard; Protein: 390 AA.
XX AC AAB66067;
XX 30-MAR-2001 (first entry)
DE Human TANGO 294 mature protein.
XX TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety.
XX Homo sapiens.
XX WO200077239-A2.
PN 21-DEC-2000.
XX 24-MAY-2000; 2000WO-US14858.
XX 14-JUN-1999; 99US-0333159.
XX (MILL-) MILLENNIUM PHARM INC.
XX McCarthy SA, Fraser CC, Sharp JD, Barnes TW;
PI WPT: 2001-032313/04.
XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT screening assays and diagnostic assays and for the treatment of
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT disease -
XX Claim 8; Pages 324-325; 359pp; English.
XX The present invention relates to TANGO or INTERCEPT proteins and coding
CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC sequences are useful for the treatment of neurological disorders such as
CC central nervous system (CNS) disorders, CNS-related disorders, focal
CC brain disorders, global-diffuse cerebral disorders and other
CC neurological and cerebrovascular disorders. The CNS disorders include
CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,

CC autonomic function disorders such as hypertension and sleep disorders,
CC neuropsychiatric disorders, psychoactive substance use disorders,
CC anxiety, and bipolar affective disorder.
XX SQ Sequence 390 AA;
Query Match 92.4%; Score 2076; DB 22; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.9e-202;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 34 VHMPTKAVDPEAFNMISEIIHQHGYPCPEEVEVATEDGYILSVNRIPRGLVQPKTGRSPV 93
Db 1 VHMPTKAVDPEAFNMISEIIHQHGYPCPEEVEVATEDGYILSVNRIPRGLVQPKTGRSPV 60
QY 94 VLLQHLVGGASNNISNLPNNSLGFIADAGFDVWGNRSGNANSRKHKTLSIDQDEFWA 153
Db 61 VLLQHLVGGASNNISNLPNNSLGFIADAGFDVWGNRSGNANSRKHKTLSIDQDEFWA 120
QY 154 FSYDEMARFDPAPVINFILQKTGOEKIYVYGYSQGTMGFIATSTMPPELAQIKMYFALA 213
Db 121 FSYDEMARFDPAPVINFILQKTGOEKIYVYGYSQGTMGFIATSTMPPELAQIKMYFALA 180
QY 214 PIATVYKHAKSPGTFFLLPDMMIKGLFGKKEFLYQTRFLRQVILDOICSNIM 273
Db 181 PIATVYKHAKSPGTFFLLPDMMIKGLFGKKEFLYQTRFLRQVILDOICSNIM 240
QY 274 LLLGGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFQWSETKNLEKCN 333
Db 241 LLLGGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFQWSETKNLEKCN 300
QY 334 QPTPVYRVRDMTPTAMWTGGQDWSNPEDVKMLLSEVTNLIYHKNPIEWAHVDFINGL 393
Db 301 QPTPVYRVRDMTPTAMWTGGQDWSNPEDVKMLLSEVTNLIYHKNPIEWAHVDFINGL 360
QY 394 DAPHRMYNEIHLMQOQETNLSQGRCEAVL 423
Db 361 DAPHRMYNEIHLMQOQETNLSQGRCEAVL 390

RESULT 4
AAU99164
ID AAU99164 standard; Protein: 398 AA.
XX AC AAU99164;
XX 24-SEP-2002 (first entry)
DE Human lysosomal acid lipase #1.
XX Human; enzyme; lysosomal acid lipase; lipid malabsorption illness;
KW cystic fibrosis; alcoholism; heart disease; heart attack;
KW Wolman disease; cholesterol ester storage disease; brain injury;
KW mood disorder; anxiety disorder; thought disorder; volition disorder;
KW sleep disorder; neurogenic disorder; myopathic disorder; COPD;
KW obesity; cancer; chronic obstructive pulmonary disease; diabetes;
KW cardiovascular disorder; Alzheimer's disease; Parkinson's disease;
KW anorexia; osteoarthritis; central nervous system disorder;
KW peripheral nervous system disorder.
XX Homo sapiens.
XX WO200236731-A2.
PN 10-MAY-2002.
XX 30-OCT-2001; 2001WO-EPI2518.
XX 31-OCT-2000; 2000US-244170P.
PR 29-MAY-2001; 2001US-293516P.
XX (FARB) BAYER AG.
XX Xiao Y;
PI

XX WPI; 2002-519248/55.
DR N-PSDB; ABK86569.
XX Novel human lysosomal acid lipase polypeptide, useful for treating
PT cancer, diabetes, obesity, chronic obstructive pulmonary disease,
PT peripheral or central nervous system disorder or cardiovascular
PT disorder.
XX
XX
PS Claim 25; Fig 2; 126pp; English.
XX
CC The invention relates to a purified human lysosomal acid lipase
CC polypeptide. Also included are the polynucleotide encoding the
CC lipase (or its fragment, derivative, allele or sequence at least 60%
CC identical to it), vectors, host cells, a reagent (e.g. an antisense
CC oligonucleotide) which binds to the lipase or polynucleotide (used for
CC detection and modulating/reducing the lipase activity) and an anti-lipase
CC antibody. The lipase and polynucleotide are useful for identifying
CC therapeutic agents that either increase or decrease the lipase activity.
CC The identified agent, the lipase and polynucleotide are useful for
CC treatment of a disease such as lipid malabsorption illness,
CC cystic fibrosis, alcoholism, heart disease, heart attack,
CC Wolman disease, cholesterol ester storage disease, brain injury,
CC mood disorder, anxiety disorder, thought disorder, volition disorder,
CC sleep disorder, neurogenic disorder, myopathic disorder, obesity, cancer,
CC chronic obstructive pulmonary disease (COPD), diabetes,
CC cardiovascular disorder, Alzheimer's disease, Parkinson's disease,
CC anorexia, osteoarthritis, a central nervous system disorder and
CC a peripheral nervous system disorder. The present sequence is the
CC human lysosomal lipase #1.
XX
SQ Sequence 398 AA;
Query Match 88.6%; Score 1991; DB 23; Length 398;
Best Local Similarity 100.0%; Pred. No. 9e-194;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 50 SEIIHQGYPCPEYEYVATEDGYILSVNRIPRGLVQPKKTSRPPVLLQHLGLVGASNWS 109
Db 25 SEIIHQGYPCPEYEYVATEDGYILSVNRIPRGLVQPKKTSRPPVLLQHLGLVGASNWS 84
QY 110 NLPNNSLGFILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAFSYDEMAREFOLPAVIN 169
Db 85 NLPNNSLGFILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAFSYDEMAREFOLPAVIN 144
QY 170 FILQKTQGEKIYYVYSGQTTMGFIATFSTMPDLAOKIKMYFALAPIATVKHAKSPGKFL 229
Db 145 FILQKTQGEKIYYVYSGQTTMGFIATFSTMPDLAOKIKMYFALAPIATVKHAKSPGKFL 204
QY 230 LLPDMIMKGLFGKKEFLYQTRFLRQLVYILCGQVILDOICSNIMLLGGFTNNMNSRA 289
Db 205 LLPDMIMKGLFGKKEFLYQTRFLRQLVYILCGQVILDOICSNIMLLGGFTNNMNSRA 264
QY 290 SYVAHTLAGTSVQNTLHWSQAVNSGELRAFDFNGSETKNEKCNQPPVRYRDMTVPT 349
Db 265 SYVAHTLAGTSVQNTLHWSQAVNSGELRAFDFNGSETKNEKCNQPPVRYRDMTVPT 324
QY 350 AMWTGGDNLSPNEDVKMLLSEVTNLIYHKNIPFWAHVDFIWLGDAPHRYNEIHLMOQ 409
Db 325 AMWTGGDNLSPNEDVKMLLSEVTNLIYHKNIPFWAHVDFIWLGDAPHRYNEIHLMOQ 384
QY 410 EETNLSQGRCEAVL 423
Db 385 EETNLSQGRCEAVL 398
RESULT 5
AAU99165
ID AAU99165 standard; Protein; 371 AA.
XX
AC AAU99165;
XX
DT 24-SEP-2002 (first entry)

XX Human lysosomal acid lipase #2.
DE
XX
XX
KW Human: enzyme; lysosomal acid lipase; lipid malabsorption illness;
KW cystic fibrosis; alcoholism; heart disease; heart attack;
KW Wolman disease; cholesterol ester storage disease; brain injury;
KW mood disorder; anxiety disorder; thought disorder; volition disorder;
KW sleep disorder; neurogenic disorder; myopathic disorder; COPD;
KW obesity; cancer; chronic obstructive pulmonary disease; diabetes;
KW cardiovascular disorder; Alzheimer's disease; Parkinson's disease;
KW anorexia; osteoarthritis; central nervous system disorder;
KW peripheral nervous system disorder.
XX
OS Homo sapiens.
XX
PN WO200236731-A2.
XX
XX 10-MAY-2002.
XX
XX 30-OCT-2001; 2001WO-EP12518.
XX
XX 31-OCT-2000; 2000US-244170P.
XX 29-MAY-2001; 2001US-293516P.
PR
PA (FARB) BAYER AG.
XX
PI Xiao Y;
XX
XX WPI; 2002-519243/55.
DR N-PSDB; ABK86579.
XX
XX Novel human lysosomal acid lipase polypeptide, useful for treating
PT cancer, diabetes, obesity, chronic obstructive pulmonary disease,
PT peripheral or central nervous system disorder or cardiovascular
PT disorder.
XX
XX Claim 25; Fig 5; 126pp; English.
XX
XX The invention relates to a purified human lysosomal acid lipase
CC polypeptide. Also included are the polynucleotide encoding the
CC lipase (or its fragment, derivative, allele or sequence at least 60%
CC identical to it), vectors, host cells, a reagent (e.g. an antisense
CC oligonucleotide) which binds to the lipase or polynucleotide (used for
CC detection and modulating/reducing the lipase activity) and an anti-lipase
CC antibody. The lipase and polynucleotide are useful for identifying
CC therapeutic agents that either increase or decrease the lipase activity.
CC The identified agent, the lipase and polynucleotide are useful for
CC treatment of a disease such as lipid malabsorption illness,
CC cystic fibrosis, alcoholism, heart disease, heart attack,
CC Wolman disease, cholesterol ester storage disease, brain injury,
CC mood disorder, anxiety disorder, thought disorder, volition disorder,
CC sleep disorder, neurogenic disorder, myopathic disorder, obesity, cancer,
CC chronic obstructive pulmonary disease (COPD), diabetes,
CC cardiovascular disorder, Alzheimer's disease, Parkinson's disease,
CC anorexia, osteoarthritis, a central nervous system disorder and
CC a peripheral nervous system disorder. The present sequence is the
CC human lysosomal lipase #2.
XX
SQ Sequence 371 AA;
Query Match 68.4%; Score 1536.5; DB 23; Length 371;
Best Local Similarity 80.9%; Pred. No. 1.5e-147;
Matches 297; Conservative 17; Mismatches 22; Indels 31; Gaps 2;
QY 50 SEIIHQGYPCPEYEYVATEDGYILSVNRIPRGLVQPKKTSRPPVLLQHLGLVGASNWS 109
Db 25 SEIIHQGYPCPEYEYVATEDGYILSVNRIPRGLVQPKKTSRPPVLLQHLGLVGASNWS 84
QY 110 NLPNNSLGFILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAFSYDEMAREFOLPAVIN 169
Db 85 NLPNNSLGFILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAFSYDEMAREFOLPAVIN 144
QY 170 FILQKTQGEKIYYVYSGQTTMGFIATFSTMPDLAOKIKMYFALAPIATVKHAKSPGKFL 229

CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder. The present sequence is a
 CC sequence used in a sequence homology alignment with the TANGO/INTERCEPT
 CC sequences of the present invention.

XX Sequence 399 AA;

Query Match 57.4%; Score 1289; DB 22; Length 399;
 Best Local Similarity 63.5%; Pred. No. 2.7e-122;
 Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 40 AVDPFAMNISEIIHQGYPCPEEVEVATEDGYILSVNRIIPRGLVQPKKTSRPPVLLQHG 99
 DB 28 AVDPETNMNVSIIISWGFSEELVETEDGYILCLNRIIPHRKKNHSDKGPVVFLLQHG 87
 QY 100 LVGGASWISNLPNNSLGFILADAGFDVWNGSRGNWSRKHKTLSIDQDEFWAFSYDEM 159
 DB 88 LLADSSNWTNLANSSILGFILADAGFDVWNGSRGNTWSRKHKTLSVSQDEFWAFSYDEM 147
 QY 160 ARFDLPVAVINFILQKQEKIYVYVSGQTMGFIASFMPPELAQIKMYFALAPIATVK 219
 DB 148 AKYDLPASINFILNKTQGEQVYVYVSGQTTIGFIASFQIPELAKRIKMPFALGPVASVA 207
 QY 220 HAKSPGTFKLLPDMIKLFGKKEFLYOTRFLQRLVYILCGOVILDOICSNIMLLGGF 279
 DB 208 FCTSPMAKLGRLPDHLIKLDFGKEFLPQSAFLKWLGTHTVTHILKELCGNLCFLCGF 267
 QY 280 NTNNMNSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAFDGMSGSETKNLEKNCQPTPVR 339
 DB 268 NERNLNRSVDVYTHSPAGTSVQNMLHWSQAVKFKQAFDNGSSAKNYFHYNQSYPT 327
 QY 340 YRVDRMTVPTAMTGGQDWLSNPEDVKMLLSEVTNLIYHKNIPEWAHVDFIWLGDAPHRM 399
 DB 328 YNVKMDLVPTAVMSGGHDVADYDYNILLTQITNLVFEHESIPWEHLDFIWLGDAPWRL 387
 QY 400 YNEIHLMOQ 409
 DB 388 YNKIILMRK 397

RESULT 8

AAO18227
 ID AAO18227 standard; Protein; 392 AA.

XX AAO18227;

DT 18-SEP-2002 (first entry)

XX Human lysosomal acid lipase related protein.

DE Human; lysosomal acid lipase; chromosome 10; lipaemic; Wolman disease;
 KW cholesteryl ester storage disease.

XX Homo sapiens.

XX US6387680-B1.

XX 14-MAY-2002.

XX 29-MAR-2001; 2001US-0820001.

XX 29-MAR-2001; 2001US-0820001.

XX (PEKE) PE CORP NY.

XX Merkulov GV, Ketchum KA, Di Francesco V, Beasley EM;

XX WPI; 2002-478445/51.

XX Isolated nucleic acids encoding a lysosomal acid lipase protein useful

PT for the prevention, diagnosis and treatment of severe infantile-onset
 PT Wolman disease and late-onset cholesteryl ester storage disease
 XX
 XX Disclosure; Column 57-60; 68pp; English.

CC The present invention provides the protein, coding and genomic sequences
 CC of a human lysosomal acid lipase. The sequences can be used in the
 CC identification of modulators of lipase activity in cells and tissues that
 CC express the lipase, particularly the severe infantile-onset Wolman
 CC disease and the milder late-onset cholesteryl ester storage disease
 CC (CESD), which are caused by mutations in different parts of the lysosomal
 CC acid lipase (LIPA) gene. The present sequence is a protein shown in the
 CC exemplification of the invention.

XX Sequence 392 AA;

Query Match 57.1%; Score 1282; DB 23; Length 392;
 Best Local Similarity 63.8%; Pred. No. 1.4e-121;
 Matches 234; Conservative 53; Mismatches 80; Indels 0; Gaps 0;

QY 40 AVDPFAMNISEIIHQGYPCPEEVEVATEDGYILSVNRIIPRGLVQPKKTSRPPVLLQHG 99
 DB 26 AVDPETNMNVSIIISWGFSEELVETEDGYILCLNRIIPHRKKNHSDKGPVVFLLQHG 85
 QY 100 LVGGASWISNLPNNSLGFILADAGFDVWNGSRGNWSRKHKTLSIDQDEFWAFSYDEM 159
 DB 86 LLADSSNWTNLANSSILGFILADAGFDVWNGSRGNTWSRKHKTLSVSQDEFWAFSYDEM 145
 QY 160 ARFDLPVAVINFILQKQEKIYVYVSGQTMGFIASFMPPELAQIKMYFALAPIATVK 219
 DB 146 AKYDLPASINFILNKTQGEQVYVYVSGQTTIGFIASFQIPELAKRIKMPFALGPVASVA 205
 QY 220 HAKSPGTFKLLPDMIKLFGKKEFLYOTRFLQRLVYILCGOVILDOICSNIMLLGGF 279
 DB 206 FCTSPMAKLGRLPDHLIKLDFGKEFLPQSAFLKWLGTHTVTHILKELCGNLCFLCGF 265
 QY 280 NTNNMNSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAFDGMSGSETKNLEKNCQPTPVR 339
 DB 266 NERNLNRSVDVYTHSPAGTSVQNMLHWSQAVKFKQAFDNGSSAKNYFHYNQSYPT 325
 QY 340 YRVDRMTVPTAMTGGQDWLSNPEDVKMLLSEVTNLIYHKNIPEWAHVDFIWLGDAPHRM 399
 DB 326 YNVKMDLVPTAVMSGGHDVADYDYNILLTQITNLVFEHESIPWEHLDFIWLGDAPWRL 385
 QY 400 YNEIHL 405
 DB 386 YNKIIL 392

RESULT 9

AAU30498

ID AAU30498 standard; Protein; 449 AA.

XX AAU30498;

XX 18-DEC-2001 (first entry)

XX Novel human secreted protein #989.

XX Human; vaccination; gene therapy; nutritional supplement;

XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;

XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX Homo sapiens.

XX WO200179449-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US08656.

XX 18-APR-2000; 2000US-0552929.

XX 26-JAN-2001; 2001US-0770160.

XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Llu C, Drmanac RT;
XX DR WPI: 2001-611725/70.
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy -
XX PS Claim 20; Page 301; 765pp; English.
XX The invention relates to novel human secreted polypeptides. The
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression or
XX physiological interactions of the polypeptide. Vectors comprising
XX the nucleic acids encoding the polypeptides and cells genetically
XX engineered to express them are also useful for producing the proteins.
XX The proteins are useful in genetic vaccination, testing and
XX therapy, and can be used as nutritional supplements. They may be used to
XX increase stem cell proliferation; to regulate haematopoiesis; and in
XX bone, cartilage, tendon and/or nerve tissue growth or regeneration;
XX immune suppression and/or stimulation; as anti-inflammatory agents; and
XX in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
XX sequences of novel human secreted proteins of the invention.

XX SQ Sequence 449 AA;

Query Match 53.0%; Score 1191.5; DB 22; Length 449;
Best Local Similarity 57.4%; Pred. No. 2.7e-112;
Matches 230; Conservative 59; Mismatches 109; Indels 3; Gaps 2;
Qy 12 SHRMEMWLLILVAYFQFNVSVMPTK--AVDPEAFNISEIIOHQGYPCPEEVEVATED 69
Db 47 SSRKMRVGLVCLVLTLSHSGSGGLTAVDPETNNVSEIISYMGFPSEEVIVETED 106
Qy 70 GYILSVNRIPLGLVQPKTGRPVVLLQHLGVGGASNWSLNPNSLGLFILDAGFDVWM 129
Db 107 GTILCLNRIPIGRKNHSDGPKPVVFLQHLGLADSSNVTNLGNSRLGLFILDADAIDWM 166
Qy 130 GNSRGNASRRKHTLSIDQDFWAFSYDEMARFDLPVAVINFLQKTQGEKIYVYVGSQGT 189
Db 167 GNTRGNTWSPKHTLSVSQDEWAFSYDEMAKYDLPASINFLNKTQGEQYVYVGHSGGT 226
Qy 190 TWGFTAFSTMPDLAQIKMYFALAPIATVKHAKSPGTGKFLLLPDMKIKGLFGKKEFLYQT 249
Db 227 TIGFTAFSOMLELAKLKNFANGPVASVAFCTSPAKLGRPLDHLIKDLFGDEEFLPOS 286
Qy 250 RFLROL-VIYLCGGVILDOICSNIMLLGGFTNNMNSRASVYAAHTLAGTSVQNILHW 308
Db 287 AFWKVAGVPHLATHYHILKELCGNLCFLICGFGNERNLNSRDVYVYTHSPAGTFVQNMXXH 346
Qy 309 SOAVNSGELRAFDMGSETKLEKCNQPTPVRYVRDMVTPTAMTGGODWLSNPDVKNL 368
Db 347 SOAVKFKQFQAFDGSACKNHYHYNQSYPTYNVMDLVPTAVXTGGHDLWDYVGVNIX 406
Qy 369 LSEVNTLYHKNIPEWAHVDFIWLGDAPHRYMNEIHLMOQ 409
Db 407 LTOITNLVHESIPWEHLDFTIWLGDAPRWLYNKIINLMRK 447

RESULT 10

AAAR37302
ID AAR37302 standard; Protein: 398 AA.

XX AAR37302;

XX 20-SEP-1993 (first entry)

XX

DE RGL precursor.
XX Rabbit gastric lipase; RGL; PRGLN2.1; fat; bioconversion;
XX hydrolysis; transesterification.
XX Oryctolagus cuniculus.
XX Key Location/Qualifiers
XX Protein 23..398
XX /note= "claim 1; page 10-11"
XX EP542629-A.
XX 19-MAY-1993.
XX 12-NOV-1992; 92EP-0403055.
XX 13-NOV-1991; 91FR-0013948.
XX (LJOJ) INST RECH JOUVEINAL.
XX Benicourt C, Blanchard C, Junien J;
XX WPI: 1993-161080/20.
XX N-PSDB: AAQ42310.
XX Rabbit gastric lipase, its precursor and their DNA - useful for
XX treating conditions linked to gastric lipase deficiency, such as
XX mucoviscidiosis and pancreatic exocrine insufficiency
XX Claim 1; Fig 7; 31pp; French.
XX RGL, opt. used with other lipases, are useful therapeutically
XX (1) to facilitate absorption of ingested fats in patients deficient
XX in endogenous gastric lipase and (2) to treat disorders caused by
XX inadequate lipase prodn. esp. mucoviscidiosis or pancreatic exocrine
XX deficiency. It can also be used (partic. when immobilised) for
XX enzymatic bioconversion, e.g. hydrolysis or transesterification.
XX SQ Sequence 398 AA;

Query Match 51.9%; Score 1166; DB 14; Length 398;
Best Local Similarity 54.6%; Pred. No. 9e-110;
Matches 219; Conservative 70; Mismatches 102; Indels 10; Gaps 3;
Qy 17 MWLLILVAYMFQFNVSVMPTKAVDPEAFNISEIIOHQGYPCPEEVEVATED 70
Db 1 MWVLFMVAALLSA-LGTHGLFGKSAPT---NPEVNMNLSQMSIWGYPSEKYEVTEDG 56
Qy 71 YILSVNRIPLGLVQPKTGRPVVLLQHLGVGGASNWSLNPNSLGLFILDAGFDVWM 130
Db 57 YILEVNRIPYKKNNGRQRPVFLQHLGLASASNWSLNPNSLGLFILDAGVGVWLG 116
Qy 131 NSRGNASRRKHTLSIDQDFWAFSYDEMARFDLPVAVINFLQKTQGEKIYVYVGSQGT 190
Db 117 NSRGNATWSRRNLYSPDSVEFWAFSEDEMAKYDLPATIDFVYKGTQGEKLVHVSQGT 176
Qy 191 MGFTAFSTMPDLAQIKMYFALAPIATVKHAKSPGTGKFLLLPDMKIKGLFGKKEFLYQTR 250
Db 177 IGFTAFSTNPKLAEIKFTYALAPVATVKYTKSLVNLKRLFIPTTFKIFGDKIFYPHNF 236
Qy 251 FLRQLVIYLCGGVILDOICSNIMLLGGFTNNMNSRASVYAAHTLAGTSVQNILHW 310
Db 237 FDQLATQVCSRETNLVICSNALFICGFDNSANLSRLDVVYVSHNPACTSVQNMLHWIQ 296
Qy 311 AVHSGELRAFDMGSETKLEKCNQPTPVRYVRDMVTPTAMTGGODWLSNPDVKNL 370
Db 297 AVKSGNFOAFNGWSPAQNVHFNQPTPTPTYNVMDLVPTAVXTGGHDLWDYVGVNIX 356
Qy 371 EVNTNLYHKNIPEWAHVDFIWLGDAPHRYMNEIHLMOQ 411
Db 357 KLSNLYHKELIPYNHLDFTIWAANAPQEVYNEIISMMAKDK 397

RESULT 11
AAB66068
ID AAB66068 standard; Protein; 221 AA.
XX
AC AAB66068;
XX
DT 30-MAR-2001 (first entry)
XX
DE Human TANGO 294 extracellular domain.
XX
KW TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety.
XX
OS Homo sapiens.
XX
PN WO200077239-A2.
XX
PD 21-DEC-2000.
XX
PF 24-MAY-2000; 2000WO-US14858.
XX
PR 14-JUN-1999; 99US-0333159.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
XX
DR WPI; 2001-032313/04.
XX
XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT screening assays, and diagnostic assays and for the treatment of
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT disease.
XX
PS Claim 8; Page 326; 359pp; English.
XX
CC The present invention relates to TANGO or INTERCEPT proteins and coding
CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC sequences are useful for the treatment of neurological disorders such as
CC central nervous system (CNS) disorders, CNS-related disorders, focal
CC brain disorders, global-diffuse cerebral disorders and other
CC neurological and cerebrovascular disorders. The CNS disorders include
CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC autonomic function disorders such as hypertension and sleep disorders,
CC neuropsychiatric disorders, psychoactive substance use disorders,
CC anxiety, and bipolar affective disorder.
XX
SQ Sequence 221 AA;
Query Match 51.8%; Score 1165; DB 22; Length 221;
Best Local Similarity 100.0%; Pred. No. 4.6e-110;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 34 VHMPTKAVDPAFMNISEIIHQGYPCPEEYEVATEDGYILSVNRPGLVQPKTGRPV 93
DB 1 VHMPTKAVDPAFMNISEIIHQGYPCPEEYEVATEDGYILSVNRPGLVQPKTGRPV 60
QY 94 VLLQHLGVGASNWSNLPNNSLGLFADAGFDVWNGNSRGNWSRKHKTLSDQDEFWA 153
DB 61 VLLQHLGVGASNWSNLPNNSLGLFADAGFDVWNGNSRGNWSRKHKTLSDQDEFWA 120
QY 154 FSYDEMARFDLPVAVINFILOKTOGKRIYVYSGQTTMGFIAPSTMPPELAQIKMYFALA 213
DB 121 FSYDEMARFDLPVAVINFILOKTOGKRIYVYSGQTTMGFIAPSTMPPELAQIKMYFALA 180
QY 214 PIATVKHAKSPGTFKLLLPDMMIKGLFGKKEFLYQTRFLRQ 254

DB 181 PIATVKHAKSPGTFKLLLPDMMIKGLFGKKEFLYQTRFLRQ 221
RESULT 12
AAP50322
ID AAP50322 standard; protein; 395 AA.
XX
AC AAP50322;
XX
DT 17-JAN-1992 (first entry)
XX
DE Rat lingual lipase protein.
XX
KW Lingual lipase; enzyme; EC-3.1.1.3; ss.
XX
OS Rattus rattus.
XX
PN GB2142337-A.
XX
PD 16-JAN-1985.
XX
PF 29-JUN-1984; 84GB-0016581.
XX
PR 01-JUL-1983; 83GB-0017989.
XX
PR 05-SEP-1983; 83GB-0023759.
XX
PA (CELL-) CELLTECH LTD.
XX
PI Carey NH, Williamson R;
XX
DR WPI; 1985-014450/03.
XX
DR N-PSDB; AAN50385.
XX
XX New lingual lipase protein for treatment of lipase deficiency - also
PT new pre-lingual lipase protein and related products
XX
PS Disclosure; Fig 5; 15pp; English.
XX
CC This protein may be expressed in a transformant host organism
CC and may be used for the treatment of lipase deficiency.
XX
SQ Sequence 395 AA;
Query Match 51.5%; Score 1156.5; DB 6; Length 395;
Best Local Similarity 54.7%; Pred. No. 8.2e-109;
Matches 216; Conservative 69; Mismatches 109; Indels 1; Gaps 1;
QY 17 MWLLILVAYMFQ-RNVNSVHMPTKAVDPAFMNISEIIHQGYPCPEEYEVATEDGYILSV 75
DB 1 MWLLILVAYMFQ-RNVNSVHMPTKAVDPAFMNISEIIHQGYPCPEEYEVATEDGYILSV 60
QY 76 NRIPRGLVQPKTGRPVVLLQHLGVGASNWSNLPNNSLGLFADAGFDVWNGNSRGN 135
DB 61 YRIPHGKNNSENIGKRPVYVLIQHLGLASATNWIANLPNNSLAFMLADAGYDVWLGNSRGN 120
QY 136 AWSRKHKTLSDQDEFWAFSYDENAREFDLPVAVINFILOKTOGKRIYVYSGQTTMGFIA 195
DB 121 TWSRKNYYSPDSVEFWAFSDEMAYDLPATINFIYQKTOGKRIYVGHQSORTIGFIA 180
QY 196 FSTMPPELAQIKMYFALAPATATVHKASPGTKFLLLPDMMIKGLFGKKEFLYQTRFLRQ 255
DB 181 FSTNPTLAKKIKTYALAPATVATVYQSPKLSIFITPFLFKLMFGKMKFLPHYFDDFL 240
QY 256 VIYLCGGVILDOICSNIMLLLGGFTNNMNSRASVVAHTLAGTSVQNIHLHWSQAVNSG 315
DB 241 GTEYCSREVLJOLCSNTLFIQCGFDKKNLNVSRFDVYLGHNPAQTSVQDFLWLAQLVRSG 300
QY 316 ELRAFDRGSEFNKLEKCNQPTPVRYRDMTPVTAMWTGGODWLSNPDVAKMLSEVTNL 375
DB 301 KFOAFNMGSPSQNMLHYNQKTPPEYDVSAMTPVPVAVWNGNDILADPODVAMLLPKLSNL 360
QY 376 IYHKNIPWAHVDFIWLGLDAPHRYNEIIHLMQOE 410

PT preventing immune system disorders (e.g. Crohn's disease), neurological disorders (e.g. Parkinson's disease), or cell proliferative disorders (e.g. cancers) -

PS Claim 48; Page 111-112; 122pp; English.

XX The present invention relates to the isolation of human lipid
 CC metabolism enzymes (LMM) designated LMM-1 to LMM-6, and the
 CC polynucleotide sequences encoding them. The LMM polypeptides,
 CC polynucleotides, agonists and antagonists are useful for diagnosing,
 CC treating or preventing disorders associated with aberrant expression
 CC of LMM, particularly immune system disorders (e.g. acquired
 CC immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,
 CC asthma or Crohn's disease), neurological disorders (e.g. epilepsy,
 CC Huntington's disease, dementia or Parkinson's disease), developmental
 CC disorders (e.g. Down's syndrome) or cell proliferative disorders
 CC (e.g. cancers including adenocarcinoma, leukaemia, lymphoma, melanoma,
 CC myeloma or sarcoma). The present sequence represents human LMM-4.

XX Sequence 399 AA;

Query Match 50.9%; Score 1143; DB 23; Length 399;

Best Local Similarity 54.2%; Pred. No. 2e-107;

Matches 213; Conservative 71; Mismatches 101; Indels 8; Gaps 2;

QY 18 WLLILVA-YMFORNVSVHMTKAVDPAPMNTSEIIHQGYPCPEEYEVATEDGYILSVN 76

DB 10 WMLLGSVMYDKKGN-----ANPEANNISQIIISYMGYPYEVYDVTTKDGYILGIY 62

QY 77 RIRPGLVOPKTSRPPVVLQHLVGASNWSNLPNNSLGFILADGDFVWNGNSRGNA 136

DB 63 RIRPGRGCPGRTAPKPAVYLQHLIASANWICNLPNNSLAFLLADSGYDVLGNSRGNT 122

QY 137 WSRKHKLSDQDEFWAFSDENARFDLPVINFILOKTGOEKIYVYVGYSGTGMGTAF 196

DB 123 WSRKHKLSPKSEIYWFASLDEMAKYDLPATINFIEIKTGKRIYVYVGHSGTGMGTAF 182

QY 197 STNPELAKIKMYFALAPIATVTRHAKSPGTKEFLPDMIMKGLFGKKEFLYQTRFLRLV 256

DB 183 STNPELAKIKIFALAPVTVYKTSQPMKLTLSRRVVKVLFSGDKMFHPTLFDQFIA 242

QY 257 IYLCGGVILQICSNIMLLGGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGE 316

DB 243 TKVCRNKLFRICSNFLFTLSGDFPQNLNMRSLDYLHSHNPAGTSVQNLHWSQAVNSGQ 302

QY 317 LRAFDMGSETKNLEKCNQPTVRYVRVDMTVPYAMTGGODWLNPNEDVKMLSEVTNLI 376

DB 303 LQAFDAGNSDONMMHFPQLTPPLYNITKMEVPTAINGGQDIVADPKDVENLLPQIANLI 362

QY 377 YHKNIPWAHVDFIWLGDAPHRYMNEIHLMQQ 409

DB 363 YKLLPHYNHVDYLGEDAPQEIYODLILMEE 395

RESULT 15

AAAR56870

XX AAR56870 standard; Protein; 379 AA.

XX AC

XX XX

DT 20-FEB-1995 (first entry)

XX XX

DE Canine gastric lipase.

XX XX

KW Canine gastric lipase; CGI; dog; fat; mucoviscidosis; enzyme;

KW blooconversion; exocrine pancreatic insufficiency.

XX XX

OS Canis familiaris.

XX XX

PN WO9413816-A.

XX XX

PD 23-JUN-1994.

XX XX

PF 16-DEC-1993; 93WO-FR01260.

XX 16-DEC-1992; 92FR-0015201.

PR (LJOU) INST RECH JOUVEINAL.

XX Benicourt C, Blanchard C, Junien J;

XX WPI; 1994-21789C/26.

XX N-PSDB; AAQ68386.

XX Recombinant canine gastric lipase and nucleic acid encoding it -

PT are used for improving absorption of ingested fat, treating

DR mucoviscidosis etc. and in enzymatic bio-conversions

XX Claim 13; Fig 9A; 52pp; French.

XX The sequence given below is the sequence of figure 9A, altered

CC according to the amendments described on page 2 of the appended

CC letter.

CC CGI is used to improve absorption of ingested fat, in healthy and

CC sick patients (e.g. having altered levels of gastric lipase); to

CC treat conditions associated with insufficiency (or lack) of lipases,

CC esp. mucoviscidosis or exocrine pancreatic insufficiency and partic.

CC where immobilised, for bioconversions, e.g. hydrolysis or

CC transesterification (other mammalian gastric lipases, or derivs.,

CC can be used in this application).

XX Sequence 379 AA;

QY 37 PTKAVDEAPFANISEITQHOGYPCPEEYEVATEDGYILSVNPRGLVQPKTGRPPVLL 96

DB 7 PT---NPEVTNISMOMITYMGYPAEYEVATEDGYILSDIPYGRKNSENIGRRPVAFL 63

QY 97 QHGLVGGASNWSNLPNNSLGFILADGDFVWNGNSRGNAWSRKHKLSDQDEFWAFS 156

DB 64 QHGLLASATNWSNLPNNSLAFILADGDFVWNGNSRGNTWARNLVYSPDSVEFWAFS 123

QY 157 DEMARFDLPVINFILOKTGOEKIYVYVGYSGTGMGTAFSTMPPELAKIKMYFALAPIA 216

DB 124 DEMAKYDLPATIDFLKKTGQDKLHYVGHSGTGMGTAFSTNPKLAKRIKTFYALAPVA 183

QY 217 TVKHAKSPGTKEFLPDMIMKGLFGKKEFLYQTRFLRO-LVIYLCGGVILQICSNIMLL 275

DB 184 TVKTEITLLNKLMLVPSFLFKLIFGNKIF-YPHHFFDQFLATEVCSRETVDLLCSNALFI 242

QY 276 LGGFNTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDMGSETKNLEKCNQ 335

DB 243 ICGEDTNLNSRLDYLHSHNPAGTSVQNLHWSQAVNSGELRAFDMGSETKNLEKCNQ 302

QY 336 TPVRYVRDMTVPYAMTGGODWLNPNEDVKMLSEVTNLIYHKNIPWAHVDFIWLGD 395

DB 303 MPYVYNTLDMHVPYIAVWNGNDLLADPHVDVLLSKLPNLIIYHKKIPPYNHLDFTIAMD 362

QY 396 PHRYMNEIHLML 407

DB 363 PQAVYNEIVSMW 374

Search completed: February 3, 2003, 13:48:54

Job time : 41 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:50:05 ; Search time 13 seconds
(without alignments)
656.579 Million cell updates/sec

Title: US-09-333-159-47
 Perfect score: 2247
 Sequence: 1 MLETLRQWIVSHRNEMWLL.....IHLMOEETNLSQRC EAVL 423

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	Published_Applications_AA :
1 :	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp :
2 :	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp :
3 :	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp :
4 :	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp :
5 :	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp :
6 :	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp :
7 :	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp :
8 :	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp :
9 :	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp :
10 :	/cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp :
11 :	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp :
12 :	/cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pcp :
13 :	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp :
14 :	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2247	100.0	423	9	US-10-042-431-47
2	2174	96.8	409	10	US-09-835-996A-21
3	2076	92.4	390	9	US-10-042-431-49
4	1991	88.6	374	10	US-09-811-825-4
5	1991	88.6	398	10	US-09-811-825-2
6	1289	57.4	399	9	US-10-042-431-41
7	1282	57.1	392	12	US-10-003-302-4
8	1165	51.8	221	9	US-10-042-431-50
9	1119	49.8	398	9	US-10-042-431-75
10	1090	48.5	398	9	US-10-056-744B-2
11	1031	45.9	395	12	US-10-003-303-2
12	783	34.8	144	9	US-10-042-431-52
13	400	17.8	127	9	US-10-056-744B-5
14	171	7.6	33	9	US-10-042-431-48
15	128	5.7	25	9	US-10-042-431-51
16	109.5	4.9	305	10	US-09-815-243-19337
17	103	4.6	222	10	US-09-815-243-5276
18	102	4.5	346	9	US-10-037-805-34
19	102	4.5	346	10	US-09-903-410-34

ALIGNMENTS

RESULT 1

US-10-042-431-47

03 10 042 431 47
: Sequence 47, Application US/10042431

; sequence 47, application US/100
; publication No. US20020182675A1: PUBLICATION NO: US20
: GENERAL INFORMATION:

APPLICANT: MCCARTHY, Sean A

APPLICANT: MCCARTHY, SEAN A
APPLICANT: BARNES, THOMAS M

APPLICANT: BARNES, THOMAS M
APPLICANT: FRASER, CHRISTOPHER C

APPLICANT: FRASER, Christopher C

```

; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES

```

; FILE REFERENCE: 10147-6U2

FILE REFERENCE: 10147 002
: CURRENT APPLICATION NUMBER: US/10/042,431

: CURRENT FILING DATE: 2001-10-25
 : CURRENT APPLICATION NUMBER: US/1

: CURRENT FILING DATE: 2001-10-23
 : PRIOR APPLICATION NUMBER: US 09/333,159

PRIOR FILING DATE: 1999-06-14

; PRIOR FILING DATE: 1999-08-14
; PRIOR APPLICATION NUMBER: US 09/578,063: PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US

;; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 79

; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: DETECTIC VER 3.1; SOFTWARE: F
; CEO ID NO. A

```

; SEQ ID NO 47
: LENGTH: 422

```

; LENGTH: 42
TYPE: ROM

```

; TYPE: PRT
; ORGANISM: Homo sapiens

```

; ORGANISM: HO
NC-10-043-131-17

Query Match	100.0%	Score 2247	DB 9	Length 423
-------------	--------	------------	------	------------

Query match 100.0%; SCORE 2247; DB 31
Best Local Similarity 100.0%; Pred. No. 4e-204:

Best Local Similarity 100.0%; PRED. NO. 4E-204;
Matches 423: Conservative 0: Mismatches 0: Indels 0: Gaps

017 1 MLEFI SPQWTVCHPBMEMWYI T I VAVMEQBNN/NSVHMDTKVAVDDREAPMNTSETTCHACQVDC 60

QY I MLETLRQWIVSHRMEMWLLILVAYMFQRNVNSVHMPTKAVDPEAFMNISEIIQHOGYPC 60

Db 1 MLETLRQWIVSHRMEMWLLILVAYMFQRNVNSVHMPTKAVDPEAFMNISEIIQHGYPC 60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Db 61 EEEVATEDGYILSVNRIPRGLVQPKKTCGRPVLLQHGGLVGGASNWI
NPNNSLGFIL 12

Qy	181	YYVYSGQGTMGFIAESTMPELAQKTKMYTALAPIATVVKHAKSPGTKFLLLPDMIMKGLF	240
Db	181	YYVYSGQGTMGFIAESTMPELAQKTKMYTALAPIATVVKHAKSPGTKFLLLPDMIMKGLF	240
Qy	241	GKKEFLYQTRFLRQLVYLICGQVILQICSNIMLLGGFNTNMNMSRASVYAHTLACT	300
Db	241	GKKEFLYQTRFLRQLVYLICGQVILQICSNIMLLGGFNTNMNMSRASVYAHTLACT	300
Qy	301	SVQNILHWSQAVNSGELRAFDMGSETKMLEKNQPPVRYRVDMTVPYAMTNGGQDWLS	360
Db	301	SVQNILHWSQAVNSGELRAFDMGSETKMLEKNQPPVRYRVDMTVPYAMTNGGQDWLS	360
Qy	361	NPEDVKMLLSEVTNLYHKNIPEWAHVDFIWLGDAPHRYNETIHLMQOEEFNLSQGRCE	420
Db	361	NPEDVKMLLSEVTNLYHKNIPEWAHVDFIWLGDAPHRYNETIHLMQOEEFNLSQGRCE	420
Qy	421	AVL 423	
Db	421	AVL 423	

RESULT 2

```

US-09-835-996A-21
: Sequence 21, Application US/098355996A
: Patent No. US20020142953A1
: GENERAL INFORMATION:
: APPLICANT: Ballinger, Dennis
: APPLICANT: Loeb, Debra
: APPLICANT: Montgomery, Julie
: APPLICANT: Tang, Y. Tom
: APPLICANT: Zhou, Ping
: APPLICANT: Goodrich, Ryle
: APPLICANT: Liu, Chenghua
: APPLICANT: Asundi, Vinod
: APPLICANT: Zhao, Qing
: APPLICANT: Wehrman, Tom
: APPLICANT: Drmanac, Radoje
: APPLICANT: Ren, Feiyao
: APPLICANT: Qian, Xiaohong
: APPLICANT: Wang, Dunrui
: TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
: FILE REFERENCE: 28110/35915A
: CURRENT APPLICATION NUMBER: US/09/835,996A
: CURRENT FILING DATE: 2001-04-16
: PRIOR APPLICATION NUMBER: US 60/197,137
: PRIOR FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: US 09/714,936
: PRIOR FILING DATE: 2000-11-17
: PRIOR APPLICATION NUMBER: US 09/667,298
: PRIOR FILING DATE: 2000-09-22
: PRIOR APPLICATION NUMBER: US 09/631,451
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: US 09/598,042
: PRIOR FILING DATE: 2000-06-20
: NUMBER OF SEQ ID NOS: 45
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 21
: LENGTH: 409
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-835-996A-21

```

RESULTS

```

US-10-042-431-49
; Sequence 49, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-602
; CURRENT APPLICATION NUMBER: US/10/042.431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-431-49

```

Query Match	92.4%;	Score 2076;	DB 9;	Length 390;
Best Local Similarity	100.0%;	Pred. No. 4.8e-188;		
Matches 390;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	34	VHMPKAVDPEAFNNISIIIOHQGYPCPEEYEVATGEGYILSVNRIPRGLVOPKKTGSRPV	93	
Db	1	VHMPKAVDPEAFNNISIIIOHQGYPCPEEYEVATGEGYILSVNRIPRGLVOPKKTGSRPV	60	
Qy	94	VLLQHGVLGGASWISNLNPNNSLGLFIADAGFDVWGNRGNAMSRKHKTLSIQDDEFWA	153	
Db	61	VLLQHGVLGGASWISNLNPNNSLGLFIADAGFDVWGNRGNAMSRKHKTLSIQDDEFWA	120	
Qy	154	FSYDEMARFDLPVAINFLQKTGGEKIIYVGYSGQTMTGFAFTSMPELAQIKIMYFALA	213	
Db	121	FSYDEMARFDLPVAINFLQKTGGEKIIYVGYSGQTMTGFAFTSMPELAQIKIMYFALA	180	
Qy	214	PIATVKHAKSPGTFFLLPDMNMIKGLFGKKEFLYQTRFLRLVLYLGCQVILDQICSNM	273	
Db	181	PIATVKHAKSPGTFFLLPDMNMIKGLFGKKEFLYQTRFLRLVLYLGCQVILDQICSNM	240	
Qy	274	LLLGEGFTNNMNMMSRASVYAAHTLAGTSVQNILHWSQAVNSGEURAFDWMGSETKNLEKCN	333	
Db	241	LLLGEGFTNNMNMMSRASVYAAHTLAGTSVQNILHWSQAVNSGEURAFDWMGSETKNLEKCN	300	
Qy	334	QPTPVRYRVFDMTVPYTMWTTGGQDWLSNPEDVKMLLSEVTNLIYHKNIPEWAHVDFIWGL	393	

Query Match	92.4%;	Score 2076;	DB 9;	Length 390;
Best Local Similarity	100.0%;	Pred. No. 4.8e-188;		
Matches 390;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	34	VHMPKAVDPEAFNNISIIIOHQGYPCPEEYEVATGEGYILSVNRIPRGLVOPKKTGSRPV	93	
Db	1	VHMPKAVDPEAFNNISIIIOHQGYPCPEEYEVATGEGYILSVNRIPRGLVOPKKTGSRPV	60	
Qy	94	VLLQHGVLGGASWISNLNPNNSLGLFIADAGFDVWGNRGNAMSRKHKTLSIQDDEFWA	153	
Db	61	VLLQHGVLGGASWISNLNPNNSLGLFIADAGFDVWGNRGNAMSRKHKTLSIQDDEFWA	120	
Qy	154	FSYDEMARFDLPVAINFLQKTGGEKIIYVGYSGQTMTGFAFTSMPELAQIKIMYFALA	213	
Db	121	FSYDEMARFDLPVAINFLQKTGGEKIIYVGYSGQTMTGFAFTSMPELAQIKIMYFALA	180	
Qy	214	PIATVKHAKSPGTFFLLPDMNMIKGLFGKKEFLYQTRFLRLVLYLGCQVILDQICSNM	273	
Db	181	PIATVKHAKSPGTFFLLPDMNMIKGLFGKKEFLYQTRFLRLVLYLGCQVILDQICSNM	240	
Qy	274	LLLGEGFTNNMNMMSRASVYAAHTLAGTSVQNILHWSQAVNSGEURAFDWMGSETKNLEKCN	333	
Db	241	LLLGEGFTNNMNMMSRASVYAAHTLAGTSVQNILHWSQAVNSGEURAFDWMGSETKNLEKCN	300	
Qy	334	QPTPVRYRVFDMTVPYTMWTTGGQDWLSNPEDVKMLLSEVTNLIYHKNIPEWAHVDFIWGL	393	

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-431-75

Query Match 49.8%; Score 1119; DB 9; Length 398;
Best Local Similarity 52.2%; Pred. No. 9.1e-98;
Matches 210; Conservative 73; Mismatches 107; Indels 12; Gaps 2;

Qy 17 MWLLTLVAYMFQNRVNSVHMTKAV-----DPEAFNISEIIHQGYPCEEYEVATED 69
Db 1 MWLLTUMASLI-----SVLGTGHLGKLPKPGSPVETWNISOMITYWYYPNEEYEVATED 55
Qy 70 GYLSVNRIPRGLVQPKTKGSRPVVLLQHLGVGGASNWSNLPNNSLGFLADAGFYVM 129
Db 56 GYILEVNRIPYKKNKSGNTGQRPVVFQHLGLASATNWSNLPNNSLAFILADAGFYWL 115
Qy 130 GNSRGNASRKHTLSIDODEFWAFSDYDEMARFDLPVAVINFLQKTQGEKIYYVGSQGT 189
Db 116 GNSRGNTWARRNLYSPDSVEFWAFSDEMAKYDLPATIDFIVKKTGOKQLHYVGSQGT 175
Qy 190 TGFATFSTMPLEAOKIKMYFALAPIATVYKHAKEPCTKELLPLDMWIKGLGKKEFLYQT 249
Db 176 TGFATFSTNPSLAKRIKTFYALAPVATVYKTSLNKLRFPQSLKFEIFGDKIFYPHN 235
Qy 250 RFLRLVLYLCOVLDQICSNIMLLGGFNTNMMNSRASVYAAHTLAGTSVONILHWS 309
Db 236 FDOFLATEVCSREMLNLLCSNLFICGFSKFNFTSRDLVYLSHNPAGTSVQNWFWT 295
Qy 310 QAVNSGELEAFDWSGETKLEKCNQPTPVRYVRDMVPTAMWTGGQDMLSNPDKMLL 369
Db 296 QAVKSGKQAVDWSQVQRHMYDQSPYVNTAMVPIAVWNGGKLLADPDQVGLLL 355
Qy 370 SEVTNLYHKNIPEVAHVDFIWLGLDAPHRMYNEIIHLMOQEE 411
Db 356 PKLNLVYHKEIPFYHLDIFWADPAQEVNDIVSMISEDK 397

RESULT 10

US-10-056-744B-2
; Sequence 2, Application US/10056744B
; Publication No. US2002019303A1
; GENERAL INFORMATION:

; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 58860, A HUMAN CHOLESTERYL ESTER
; TITLE OF INVENTION: HYDROLASE AND USES THEREFOR
; FILE REFERENCE: MPI2001-026PIRNM
; CURRENT APPLICATION NUMBER: US/10/056.744B
; PRIOR FILING DATE: 2002-08-16
; PRIOR FILING DATE: 60/264,167
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 398

; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-056-744B-2

Query Match 48.5%; Score 1090; DB 9; Length 398;
Best Local Similarity 52.4%; Pred. No. 5e-95;
Matches 207; Conservative 61; Mismatches 123; Indels 4; Gaps 2;

Qy 17 MWLLTLVAYMFQNRVNS---VHMPTKAVDPEAFMNISEIIHQGYPCEEYEVATEDGYIL 73
Db 2 MWLLTTTCLICGTLNAGGFLDENE-VNPEVMMNTSEIIYNGYPSEYEVTEDGYIL 60
Qy 74 SVNRIPRGLVQPKTKGSRPVVLLQHLGVGGASNWSNLPNNSLGFLADAGFYVMGNSR 133
Db 61 LVNRIPYGRTHARSTGPRPVVYMQHALFADNAYWLENYANGSLGFLADAGFYVMGNSR 120

Qy 134 GNAWSRKHTLSIDODEFWAFSDYDEMARFDLPVAVINFLQKTQGEKIYYVGSQGTMGF 193
Db 121 GNTWSRRKHTLSETDEKFWAFSDEMAKYDLPVGVDFIVNKTQGEKLYFIGHSLGTTIGF 180
Qy 194 IAFSTMPLEAOKIKMYFALAPIATVYKHAKEPCTKELLPLDMWIKGLGKKEFLYQTRFLR 253
Db 181 VAFSTMPLEAOKIKMYFALAPIATVYKHAKEPCTKELLPLDMWIKGLGKKEFLYQTRFLR 240
Qy 254 QLVYLYLCOVLDQICSNIMLLGGFNTNMMNSRASVYAAHTLAGTSVONILHWSQAVN 313
Db 241 IASTKICNKKILWLJCSEFMSLWAGSNKKNNQSRMDVYMSHAPTSSVHNILHIKQLYH 300
Qy 314 SGELEAFDWSGETKLEKCNQPTPVRYVRDMVPTAMWTGGQDMLSNPDKMLLSEVT 373
Db 301 SDEFRAYDNGDADNKKHYNQSHPPYIDLTAKVKPTAIWAGGHVLTVPDVARILPQIK 360
Qy 374 NLYHKNIPEVAHVDFIWLGLDAPHRMYNEIIHLMO 408
Db 361 SLHYFKLLPDWNRHFDVWGLDAPQRMYSIIALMK 395

RESULT 11

US-10-003-302-2
; Sequence 2, Application US/10003302
; Patent No. US20020142435A1
; GENERAL INFORMATION:

; APPLICANT: MERKULOV, Gennady et al.
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01186DIV
; CURRENT APPLICATION NUMBER: US/10/003.302
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 395

; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-003-302-2

Query Match 45.9%; Score 1031; DB 12; Length 395;
Best Local Similarity 51.0%; Pred. No. 1.8e-89;
Matches 199; Conservative 61; Mismatches 124; Indels 6; Gaps 4;

Qy 17 MWLLTLVAYMFQNRVNS---VHMPTKAVDPEAFMNISEIIHQGYPCEEYEVATEDGYIL 73
Db 2 MWLLTTTCLICGTLNAGGFLDENE-VNPEVMMNTSEIIYNGYPSEYEVTEDGYIL 60
Qy 74 SVNRIPRGLVQPKTKGSRPVVLLQHLGVGGASNWSNLPNNSLGFLADAGFYVMGNSR 133
Db 61 LVNRIPYGRTHARSTGPRPVVYMQHALFADNAYWLENYANGSLGFLADAGFYVMGNSR 120
Qy 134 GNAWSRKHTLSIDODEFWAFSDYDEMARFDLPVAVINFLQKTQGEKIYYVGSQGTMGF 193
Db 121 GNTWSRRKHTLSETDEKFWAFSDEMAKYDLPVGVDFIVNKTQGEKLYFIGHSLGTTIGF 180
Qy 194 IAFSTMPLEAOKIKMYFALAPIATVYKHAKEPCTKELLPLDMWIKGLGKKEFLYQTRFLR 253
Db 181 VAFSTMPLEAOKIKMYFALAPIATVYKHAKEPCTKELLPLDMWIKGLGKKEFLYQTRFLR 240
Qy 254 QLVYLYLCOVLDQICSNIMLLGGFNTNMMNSRASVYAAHTLAGTSVONILHWSQAVN 313
Db 241 IASTKICNKKILWLJCSEFMSLWAGSNKKNNQSRMDVYMSHAPTSSVHNILHIKQLYH 300
Qy 314 SGELEAFDWSGETKLEKCNQPTPVRYVRDMVPTAMWTGGQDMLSNPDKMLLSEVT 373
Db 301 SDEFRAYDNGDADNKKHYNQSHPPYIDLTAKVKPTAIWAGGHVLTVPDVARILPQIK 360
Qy 374 NLYHKNIPEVAHV-DFIWLGLDAPHRMYN 401
Db 361 SLSLSLSLPEWEPTDFEYVWGLDAPQRMFS 390

RESULT 12

US-10-042-431-52
; Sequence 52, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:

; APPLICANT: MCCARTHY, Sean A

; APPLICANT: BARNES, Thomas M

; APPLICANT: FRASER, Christopher C

; APPLICANT: SHARP, John D

; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES

; FILE REFERENCE: 10147-602

; CURRENT APPLICATION NUMBER: US/10/042,431

; CURRENT FILING DATE: 2001-10-25

; PRIOR APPLICATION NUMBER: US 09/333,159

; PRIOR FILING DATE: 1999-06-14

; PRIOR APPLICATION NUMBER: US 09/578,063

; PRIOR FILING DATE: 2000-05-24

; NUMBER OF SEQ ID NOS: 79

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 52

; LENGTH: 144

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-042-431-52

Query Match

Best Local Similarity 34.8%; Score 783; DB 9; Length 144;

Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 NTNNMNSRASYAAHTLAGTSVQNLHWSQAVNSGELRAFDWGS TKNLKCNQPTPVR 339

DB 1 NTNNMNSRASYAAHTLAGTSVQNLHWSQAVNSGELRAFDWGS TKNLKCNQPTPVR 60

QY 340 YVRDNTVPTAMTGGQDLNPNEDVKMLLSEVTNLIYHKNIPEWAHVDFINGLDAPHRM 399

DB 61 YVRDNTVPTAMTGGQDLNPNEDVKMLLSEVTNLIYHKNIPEWAHVDFINGLDAPHRM 120

QY 400 YNEIHLMOQEEETNLSQGRCEAVL 423

DB 121 YNEIHLMOQEEETNLSQGRCEAVL 144

QY 400 YNEIHLMOQEEETNLSQGRCEAVL 423

DB 121 YNEIHLMOQEEETNLSQGRCEAVL 144

RESULT 13

US-10-056-744B-5
; Sequence 5, Application US/10056744B
; Publication No. US20020193303A1
; GENERAL INFORMATION:

; APPLICANT: Kapeller-Libermann, Rosana

; TITLE OF INVENTION: 58860, A HUMAN CHOLESTERYL ESTER

; TITLE OF INVENTION: HYDROLASE AND USES THEREFOR

; FILE REFERENCE: MPI2001-026PIRNM

; CURRENT APPLICATION NUMBER: US/10/056,744B

; CURRENT FILING DATE: 2002-08-16

; PRIOR APPLICATION NUMBER: 60/264,167

; PRIOR FILING DATE: 2001-01-25

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 127

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-056-744B-5

Query Match

Best Local Similarity 17.8%; Score 400; DB 9; Length 127;

Matches 72; Conservative 22; Mismatches 29; Indels 0; Gaps 0;

QY 107 WISNLPNNSLIGFLADAGFDVWNGNSRKNKHTLTSDQDEFWAFSYDENAREDLPA 166

DB 1 WTANGPNKSLAFILADAGIDYDVLGNRGNTYSRKHVKLNPSHSEFWDFSWHGMGYDLPA 60

QY 107 WISNLPNNSLIGFLADAGFDVWNGNSRKNKHTLTSDQDEFWAFSYDENAREDLPA 166

DB 1 WTANGPNKSLAFILADAGIDYDVLGNRGNTYSRKHVKLNPSHSEFWDFSWHGMGYDLPA 60

QY 167 VINFILOKTOGEKIYYGVYSGQTTMGFIATFSTMPELAOKIKMYFALAPIATVYKHAKSPGT 226
DB 61 MIDYVLETTGQEKLVGHSGSQTTTFVFMVLSRPEVNEKIKSHALAPVAYWKNHVRSLV 120

QY 227 KFL 229

DB 121 KLL 123

RESULT 14

US-10-042-431-48

; Sequence 48, Application US/10042431

; Publication No. US20020182675A1

; GENERAL INFORMATION:

; APPLICANT: MCCARTHY, Sean A

; APPLICANT: BARNES, Thomas M

; APPLICANT: FRASER, Christopher C

; APPLICANT: SHARP, John D

; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,

; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES

; FILE REFERENCE: 10147-602

; CURRENT APPLICATION NUMBER: US/10/042,431

; CURRENT FILING DATE: 2001-10-25

; PRIOR APPLICATION NUMBER: US 09/333,159

; PRIOR FILING DATE: 1999-06-14

; PRIOR APPLICATION NUMBER: US 09/578,063

; PRIOR FILING DATE: 2000-05-24

; NUMBER OF SEQ ID NOS: 79

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 48

; LENGTH: 33

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-042-431-48

Query Match

Best Local Similarity 7.6%; Score 171; DB 9; Length 33;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLETLSRQWIVSHRMEMWLLILVAYMFQNRVNS 33

DB 1 MLETLSRQWIVSHRMEMWLLILVAYMFQNRVNS 33

RESULT 15

US-10-042-431-51

; Sequence 51, Application US/10042431

; Publication No. US20020182675A1

; GENERAL INFORMATION:

; APPLICANT: MCCARTHY, Sean A

; APPLICANT: BARNES, Thomas M

; APPLICANT: FRASER, Christopher C

; APPLICANT: SHARP, John D

; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,

; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES

; FILE REFERENCE: 10147-602

; CURRENT APPLICATION NUMBER: US/10/042,431

; CURRENT FILING DATE: 2001-10-25

; PRIOR APPLICATION NUMBER: US 09/333,159

; PRIOR FILING DATE: 1999-06-14

; PRIOR APPLICATION NUMBER: US 09/578,063

; PRIOR FILING DATE: 2000-05-24

; NUMBER OF SEQ ID NOS: 79

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 51

; LENGTH: 25

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-042-431-51

Query Match

Best Local Similarity 5.7%; Score 128; DB 9; Length 25;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: February 3, 2003, 13:54:11
Job time : 14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:47:54 ; Search time 21 Seconds
(without alignments)
1936.422 Million cell updates/sec

Title: US-09-333-159-47

Perfect score: 2247

Sequence: 1 MLETLRQWIVSHRMEMWLL.....IHLMOQETNLSQRCRAVL 423

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR.73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	1290	57.4	399	2 S41408	lysosomal acid lip
2	1290	57.4	399	2 G01416	lysosomal acid lip
3	1164.5	51.8	395	1 L1R7T	triacylglycerol li
4	1139	50.7	397	1 JC4017	triacylglycerol li
5	1119	49.8	398	2 S07145	triacylglycerol li
6	752.5	33.5	403	2 T33198	hypothetical prote
7	751	33.4	411	2 T22290	hypothetical prote
8	730	32.5	405	2 T22675	hypothetical prote
9	707.5	31.5	405	2 H88930	protein R1G1.14
10	692	30.8	411	2 C89074	protein K048.5 [1
11	655.5	29.2	426	2 T20480	hypothetical prote
12	574.5	25.6	559	2 JT0949	egg-specific prote
13	505	22.5	443	2 T39540	triacylglyceride lipas
14	478.5	21.3	1585	2 T31611	hypothetical prote
15	475	21.1	467	2 T41053	triacylglyceride lipas
16	445	19.8	344	2 E84526	probable lysosomal
17	422.5	18.8	548	2 S37969	probable triacylgl
18	412.5	18.4	460	2 T39443	probable triglycer
19	411.5	18.3	413	2 T43170	probable triacylgl
20	294	13.1	431	2 D86318	protein F15H18.6 [
21	285	12.7	538	2 S64842	probable membrane
22	271.5	12.1	573	2 S64754	probable membrane
23	234.5	10.4	509	2 G96766	protein lipase F2P
24	149	6.6	62	2 S59904	lipase, esterase
25	143	6.4	336	2 F83425	probable pregastric
26	126	5.6	987	2 A64474	hypothetical prote
27	113	5.0	412	2 B72391	conserved hypothet
28	112.5	5.0	40	2 S19539	triacylglycerol li
29	107	4.8	281	2 A11912	lysoospholipase

ALIGNMENTS

RESULT 1

S41408

lysosomal acid lipase (EC 3.1.1.1.-) / sterol esterase (EC 3.1.1.13) precursor - hum
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
C:Accession: S41408; A39315; S47187
R:Ames, D.; Merkel, M.; Eckerskorn, C.; Greten, H.
Eur. J. Biochem. 219, 905-914, 1994
A:Title: Purification, characterization and molecular cloning of human hepatic lys.
A:Reference number: S41408; MUID:94155897; PMID:8112342
A:Accession: S41408
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-399 <AME>
A:Cross-references: EMBL:X76488; NID:g434305; PIDN:CAA54026.1; PID:g434306
J. Biol. Chem. 265, 22479-22484, 1991
A:Title: Cloning and expression of cDNA encoding human lysosomal acid lipase/chole
A:Reference number: A39315; MUID:92042192; PMID:1718995
A:Accession: A39315
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-15, 'P', 17-399 <AND>
A:Cross-references: GB:M74775; NID:g187151; PIDN:AAA59519.1; PID:g187152
R:Du, H.; Gregory, G.A.
submitted to the EMBL Data Library, April 1994
A:Description: Structural conservation of putative functional motifs between mouse

hypothetical prote
hypothetical prote
probable protease
OutR protein - Eme
platelet-activatin
probable 2-acetyl-
hypothetical prote
probable permease
ABC transporter, m
unknown protein F2
probable epha prot
hypothetical prote
hypothetical prote
reverse transcript
probable membrane
DNA polymerase I V

Query Match 57.4%; Score 1290; DB 2; Length 399;
Best Local Similarity 61.0%; Pred. No. 1.8e-105;
Matches 242; Conservative 59; Mismatches 94; Indels 2; Gaps 1;

Qy 15 MENWLLIYVAFQFNVSVMPTK--AVDPAFNISIIHQHQPCEYEVAEDGVY 72
Db 1 MKNRFLGLVWCLVLTLSHSGSGKLTAVDPETNNVSEIISYWGFPPEEYLVETEDGVY 60
Qy 73 LSVNRIPLGLVQPKTGRPVVLLQHLGVGGASNIPLNPNNSLGLADAGFDVWGMNS 132
Db 61 LCLINRPHGRKNHSDKPKVYVFLQRLGLADSSNVTNLANSLSGLADAGFDVWGMNS 120

QY 133 RGNWSRKHKTLSIDODEFWAFSYDEMAREDLPAVINFILOKTGOEKIYVYVSGQTTMG 192
Db 121 RGNWSRKHKTLSVSDQEFWAFSYDEMAKYDLPASINFILOKTGOEQVYVGHSGQTTIG 180
QY 193 FIAPSTMPDLAQIKMYFALAPATATVKAHSGPTKFLLLPDMIMKGLFGKKEFLYQTRFL 252
Db 181 FIAPSQIPDLAKRIMKMFALGPVAVAFCSFPAKGLRGLPDHLIKLDFGDKREFLPQSAFL 240
QY 253 ROLVYLCGVILDOICSNMILLGGFTNTNMNMSRVSVAHAHTLAGTSVQNILHWSQAV 312
Db 241 KWLGTHTVCTHVKELCGNLCFLCGGNERNLNMSRDVVTTHSPAGTSVQNMLHWSQAV 300
QY 313 NSGELRAFDMGSETKNEKCNQPTPVRYRDMTPVPTAMTGGODWLSNPDYKMLLSEV 372
Db 301 KFQFOAFDMGSSAKNFHYNQSPPTTYNKDMLVPTAVMSGGHDLADYVDVNNILTOI 360
QY 373 TNLVYHKNIPFWHVDPIWGLDAPHRYNEIHLMOQ 409
Db 361 TNLVHESIPWEHLDFIWGLDAPWRLYNKIINLMRK 397

RESULT 2
G01416

lysosomal acid lipase - human

C:Species: Homo sapiens (man)

C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 18-Jun-1999

C:Accession: G01416

R:Du, H.

submitted to the EMBL Data Library, April 1994

A:Reference number: G06919

A:Accession: G01416

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-399 <DUX>

A:Cross-references: EMBL:U08464; NID:G505052; PIDN:AAB60328.1; PID:G505053

C:Superfamily: triacylglycerol lipase, lingual

C:Keywords: glycoprotein

Query Match 57.4%; Score 1290; DB 2; Length 399;
Best Local Similarity 61.0%; Pred. No. 1.8e-105;
Matches 242; Conservative 59; Mismatches 94; Indels 2; Gaps 1;

QY 15 MEMLLILVAYMFQNVNVMPTK--AVDPAFPMNISEIIHQGYPCPEEYEVATEDGYI 72
Db 1 MKRFLGLVCLVLTLSHSGSRGKLAVDPTNMYNYSIIISYNGFPSEYLVETEDGYI 60
QY 73 LSVNRIPLGLVQPKTSRPVLLQHLGVLGASNWSINLNNLSGLFLADAGFDVWNGNS 132
Db 61 LCLNRIPLHGRKNSHSDGKPKPVYFLQHLGLADSSNNWVNLANSLSGLFLADAGFDVWNGNS 120
QY 133 RGNWSRKHKTLSIDODEFWAFSYDEMAREDLPAVINFILOKTGOEKIYVYVSGQTTMG 192
Db 121 RGNWSRKHKTLSVSDQEFWAFSYDEMAKYDLPASINFILOKTGOEQVYVGHSGQTTIG 180

QY 193 FIAPSTMPDLAQIKMYFALAPATATVKAHSGPTKFLLLPDMIMKGLFGKKEFLYQTRFL 252
Db 181 FIAPSQIPDLAKRIMKMFALGPVAVAFCSFPAKGLRGLPDHLIKLDFGDKREFLPQSAFL 240
QY 253 ROLVYLCGVILDOICSNMILLGGFTNTNMNMSRVSVAHAHTLAGTSVQNILHWSQAV 312
Db 241 KWLGTHTVCTHVKELCGNLCFLCGGNERNLNMSRDVVTTHSPAGTSVQNMLHWSQAV 300
QY 313 NSGELRAFDMGSETKNEKCNQPTPVRYRDMTPVPTAMTGGODWLSNPDYKMLLSEV 372
Db 301 KFQFOAFDMGSSAKNFHYNQSPPTTYNKDMLVPTAVMSGGHDLADYVDVNNILTOI 360
QY 373 TNLVYHKNIPFWHVDPIWGLDAPHRYNEIHLMOQ 409
Db 361 TNLVHESIPWEHLDFIWGLDAPWRLYNKIINLMRK 397

RESULT 3
L1RTT

triacylglycerol lipase (EC 3.1.1.3) precursor, lingual - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 11-Jun-1999
C:Accession: A23045
R:Doeherty, A.J.P.; Bodmer, M.W.; Angal, S.; Verger, R.; Riviere, C.; Love, P.A.; Lyo
Nucleic Acids Res. 13, 1891-1903, 1985
A:Title: Molecular cloning and nucleotide sequence of rat lingual lipase cDNA.
A:Reference number: A23045; MUID:85215587; PMID:3839077
A:Accession: A23045
A:Molecule type: mRNA
A:Residues: 1-395 <DUC>

A:Cross-references: GB:X02309; NID:G56595; PIDN:CAA26179.1; PID:G56596

A:Experimental source: strain Sprague-Dawley

A:Note: The partial sequence of the mature protein from a different, unspecified stra

C:Comment: This acid-stable lipase is secreted by the serous (von Ebner's) glands at

C:Superfamily: triacylglycerol lipase, lingual

C:Keywords: carboxylic ester hydrolase; glycoprotein; lipid digestion; saliva; serous

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F:33,68,98,184,270/Binding site: carbohydrate (Asn) (covalent) #status Predicted

Query Match 51.8%; Score 1164.5; DB 1; Length 395;
Best Local Similarity 54.9%; Pred. No. 1.9e-94;

Matches 217; Conservative 69; Mismatches 108; Indels 1; Gaps 1;

QY 17 MWLLILVAYNFQ-RNVNSVHMPKAVDPAFPMNISEIIHQGYPCPEEYEVATEDGYILSV 75

Db 1 MWLLITSVISTFGAGHLGKLGCPGNPEANMISQMTYWGYPQCEYEVATEDGYILGV 60

QY 76 NRIPRGLVQPKTSRPVLLQHLGVLGASNWSINLNNLSGLFLADAGFDVWNGNSRGN 135

Db 61 YRIPHGKNNSENIKGRPVYVQLHGLIASATNWTANLNNLSLAFLADAGFDVWNGNSRGN 120

QY 136 AWSRKHKTLSIDODEFWAFSYDEMAREDLPAVINFILOKTGOEKIYVYVSGQTTMGFIA 195

Db 121 TWSRKNVYSPDSVEFWAFSDEMAREDLPAVINFILOKTGOEKIYVYVSGQTTMGFIA 180

QY 196 FSTMPDLAQIKMYFALAPATATVKAHSGPTKFLLLPDMIMKGLFGKKEFLYQTRFLRQL 255

Db 181 FSTNPTLAKKIKFYALAPATATVKAHSGPTKFLLLPDMIMKGLFGKKEFLYQTRFLRQL 240

QY 256 VIYLCGVILDOICSNMILLGGFTNTNMNMSRVSVAHAHTLAGTSVQNILHWSQAVNSG 315

Db 241 GTEVCSREVLDDLCNTLFIKCGFDKKNLNVSRPDVYLGHNPAGTSVQDFLHMAQLVRS 300

QY 316 ELRAFDMGSETKNEKCNQPTPVRYRDMTPVPTAMTGGODWLSNPDYKMLLSEVTLN 375

Db 301 KFOAFNMGSSFSQNNMLHYNQKTPPEYDVSAMTPVYVWNGNDILADPDVAMLLPKLSNL 360

QY 376 IYHKNIPFWHVDPIWGLDAPHRYNEIHLMOQ 410

Db 361 LFHKEILAYNHLDFIWADAPQEVYNEISMMAED 395

RESULT 4

JC4017

triacylglycerol lipase (EC 3.1.1.3) PGE precursor - bovine

N:Alternate names: pregastric esterase

C:Species: Bos primigenius taurus (cattle)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: JC4017; S64678

R:Timmermans, M.Y.J.; Teuchy, H.; Kupers, L.P.M.

Gene 147, 259-262, 1994

A:Title: the cDNA sequence encoding bovine pregastric esterase.

A:Reference number: JC4017; MUID:95011625; PMID:7926811

A:Accession: JC4017

A:Molecule type: mRNA

A:Residues: 1-397 <TIM>

A:Cross-references: GB:L26319; NID:G600756; PIDN:AAA57037.1; PID:G600757

A:Experimental source: tongue

R:Timmermans, M.Y.J.; Reekmans, G.; Teuchy, H.J.H.; Kupers, L.P.M.

Biochem. J. 314, 931-936, 1996

A:Title: Inhibition studies on calf pregastric esterase: the enzyme has no functional

A:Reference number: S64678; MUID:96177869; PMID:8615791

A:Accession: S64678

A:Molecule type: protein

A:Residues: 24-40:248-253 <NID>

C:Comment: Pregastric esterase is a major fat-digesting enzyme.

C:Genetics:

A:Gene: pge

C:Superfamily: triacylglycerol lipase, lingual

C:Keywords: blocked amino end; carboxylic ester hydrolase; glycoprotein

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-397/Product: pregastric esterase #status predicted <MAT>

F:33,270,326/Binding site: carboxylate (Asn) (covalent) #status predicted

F:171/Active site: Ser #status predicted

Query Match 50.7%; Score 1139; DB 1; Length 397;
Best Local Similarity 53.7%; Pred. NO. 3.4e-92;
Matches 212; Conservative 68; Mismatches 113; Indels 2; Gaps 2;

QY 18 WLLILVAYMFORNVNSVHPTKAVDPEAFNMNISEIIHQGYPCPEYEVATEDGYILSVNR 77

Db 3 WLVTVCFTHMSGNAFCFLGKIAPKPEASMNYSOMISYMGVPSSEMHKVTADGYILOVYR 62

QY 78 IPRGLVQPKTKSRPVVLLQHLGVLGASNWSNLSNLSGFTLADAGDVMNGSRGNW 137

Db 63 IPRGNANHLQCRPVVFLQHLGASATNWSNLSNLSGFTLADAGDVMNGSRGNW 122

QY 138 SRKHKTSLSDODEWAFSDYDEMAFDPALPAVINFILOKTGOEKIYVYVSGQTMGFIAPS 197

Db 123 AQEHLVYSPDSPEFWAFSEDEMAEYDLPSTIDFLLRRTGOKKLHYVGHSGQTTIGFIAPS 182

QY 198 TMPELAOKTKMYFALAPATVYKHAASPGTKFLLPDMIMKGLFGKKEFLYOTRFRLQ-IV 256

Db 183 TSPTLAERIKYFALAPATVYKHAASPGTKFLLPDMIMKGLFGKKEFLYOTRFRLQ-IV 241

QY 257 IYLCGOVILDOCSNMLLGGFTNNMNSRASYAHTLAGTSVQNLHLSQAVNSGE 316

Db 242 VECRSRETLVLCCKNALFAITGVGNKFNMSRLDYIAHNPAGTSVQNLHLSQAVNSGE 301

QY 317 LRAFDGSGETKLEKNCQPTPVYRVDMTPTVMTGQDWSNPEDVKMLLSEVTNLI 376

Db 302 FOAFDAGPYQNLHMHVQPTPIYNTLANNVPIAVNSADNLLADPDQVDVFLSLKSLNLI 361

QY 377 YKKNIDPEAHVDFIWLGDAPHRYNEIHLMOQEE 411

Db 362 YKKEIPNHLDFIWMADAPQEVYNEIVSLMAEDK 396

RESULT 5

S07145

triacylglycerol lipase (EC 3.1.1.3) precursor, gastric - human

C:Species: Homo sapiens (man)

C:Date: 29-Jan-1993 #sequence revision 29-Jan-1993 #text_change 18-Jun-1999

C:Accession: S07145; S27102; S04942

R:Rodmer, M.W.; Angal, S.; Varranton, G.T.; Harris, T.J.R.; Lyons, A.; King, D.J.; Ptero

Biochim. Biophys. Acta 909, 237-244, 1987

A:Title: Molecular cloning of a human gastric lipase and expression of the enzyme in yea

A:Reference number: S07145; MUID:87299724; PMID:3304425

A:Accession: S07145

A:Molecule type: mRNA

A:Residues: 1-398 <R0DI>

A:Cross-references: EMBL:X05997; NID:g31771; PIDN:CAA29413.1; PID:g758063

A:Accession: S27102

A:Molecule type: protein

A:Residues: 20-24,'X',26,'X',28-33,'X',35-42,'X',44,'X',46-48,'X',50,'X',52,'X',54-56 <E

A:Note: It is uncertain whether Met-1 or Met-7 is the initiator

R:Bernbaeck, S.; Blaeckberg, L.

Eur. J. Biochem. 182, 495-499, 1989

A:Title: Human gastric lipase. The N-terminal tetrapeptide is essential for lipid bindin

A:Reference number: S04942; MUID:89325292; PMID:2753032

A:Accession: S04942

A:Molecule type: protein

A:Residues: 20-45 <BER>

C:Superfamily: triacylglycerol lipase, lingual

C:Keywords: carboxylic ester hydrolase; extracellular protein; glycoprotein; lipid
F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-398/Product: triacylglycerol lipase, gastric #status experimental <MAT>

F:34,99,185,271/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 49.8%; Score 1119; DB 2; Length 398;

Best Local Similarity 52.2%; Pred. NO. 2e-90;

Matches 210; Conservative 73; Mismatches 107; Indels 12; Gaps 2;

QY 17 MWLLLVAYMFORNVNSVHPTKAV-----DPEAFNMNISEIIHQGYPCPEYEVATED 69

Db 1 MWLLLVAYMFORNVNSVHPTKAV-----SVLGTTHGLFGKLPQSPVTHNLSNLSGFTLADAGDVM 55

QY 70 GYILSVNRIPRGLVQPKTKSRPVVLLQHLGVLGASNWSNLSNLSGFTLADAGDVM 129

Db 56 GYILEVNRIPYKKNKSGTQRPVFLQHLGASATNWSNLSNLSGFTLADAGDVM 115

QY 130 GNSRGNASRKHKTSLSDODEWAFSDYDEMAFDPALPAVINFILOKTGOEKIYVYVSGQT 189

Db 116 GNSRGTWARRNLYYSPDSPEFWAFSEDEMAKYDLPATIDFIVKTKGQKLHYVGHSGQT 175

QY 190 TMGFTAFSTMPLELAOKIKMYFALAPATVYKHAASPGTKFLLPDMIMKGLFGKKEFLYOT 249

Db 176 TIGFTAFSTNLSARIKTFYALAPATVYKHAASPGTKFLLPDMIMKGLFGKKEFLYOT 235

QY 250 RFLRLVLYLCGOVILDOCSNMLLGGFTNNMNSRASYAHTLAGTSVQNLHLS 309

Db 236 FFIQFLATEVCSREMLNLLCSNALFICGFSKFNFTSLDVLVLSHPAGTSVQNMFWHT 295

QY 310 QAVNSGELRAFDMGSETKLEKNCQPTPVYRVDMTPTVMTGQDWSNPEDVKMLL 369

Db 296 QAVKSGKQAYDMGSPVQNRHMDQSPPYNYNTAMNVPVAVNMGKDLLADPDQVGLLL 355

QY 370 SEVTNLIYKNIPEAHVDFIWLGDAPHRYNEIHLMOQEE 411

Db 356 PKLPNLIYKHEIPNHLDFIWMADAPQEVYNDIVSMISEDK 397

RESULT 6

T33198

hypothetical protein ZK6.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence revision 29-Oct-1999 #text_change 21-Jan-2000

C:Accession: T33198

R:Wu, X.

submitted to the EMBL Data Library, May 1998

A:Description: The sequence of C. elegans cosmid ZK6.

A:Reference number: Z21301

A:Accession: T33198

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-403 <WUX>

A:Cross-references: EMBL:AF067942; PIDN:AAC17694.1; GSPDB:GN00023; CESP:ZK6.7

A:Experimental source: strain Bristol N2; clone ZK6

C:Genetics:

A:Gene: CESP:ZK6.7

A:Map position: 5

A:Introns: 23/3; 63/1; 219/3; 319/3; 365/3

C:Superfamily: triacylglycerol lipase, lingual

Query Match

Best Local Similarity 33.5%; Score 752.5; DB 2; Length 403;

Matches 168; Conservative 66; Mismatches 142; Indels 33; Gaps 12;

QY 17 MW-LLILVAYMFORNVNSVHPTKAVDPEAFNMNISEIIHQGYPCPEYEVATEDGYILSV 75

Db 1 MRFVFLTAFAFFVGVGSH-----GDPELHMTTPQIERWGYPMIYTVATDDGYILEM 55

QY 76 NRIPRGLVQPKTK-----GSRPVVLLQHLGVLGASNWSNLSNLSGFTLADAGDVM 129

Db 56 HRIIPG-----KTNVTWPNKRPVVFQVHGLICASSDWWVNLDPQAGFLFADAGDVM 110

QY 130 GNSRGNASRKHKTSLSDODEWAFSDYDEMAFDPALPAVINFILOKTGOEKIYVYVSGQT 189

```

111 GNRGNTYSMKHDKLSPSHAFDWSDEMATYDLNAMINHVLEVTQDSVYVYMGHSGQT 170
190 TMGFIAPSTWP-ELAAKIKMYFALAPATYKHAQSPGTEKLLLPDMNMGK--LFGKKEF 245
171 LWFHSLSKDGSFAKAKKEFFALAPIGSVKHLKGLFUSFFANTFSLFEDGWEDLFGAGEF 230
246 LYQTRFLRQLVILYCGOVIDL-QICSNMILLGSGFNTNNMNSRASVYAAHTLAGTSVQN 304
231 LPNWMKLAADICGKGLVEADLCNVLFLIAGPESDQWNTVRPVYATHDPAGTSTQN 290
305 ILHWSQAVNSGELRAFDMGSETKNLEKCNQPTPVYRVYRDMT-----VPTAMWTGGQDWLS 360
291 IVHWMQVHGGVPAYDMGKT-NKKYGOANPEY---DFTAIKGTKIYLYNSDADWLA 346
361 NPEDV-KMLLSEVTLNLYKHN-IPEWAHVDFTWGLDAPHRMYNEIHL 406
347 DTPDVPDYLTRLNPAIVAQNHLPLDYNHLDFTWGLRAPDODIYRPAIKL 395

RESULT 7
T22290
hypothetical protein F46B6.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T22290
R:McMurray, A.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19542
A:Accession: T22290
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-411<N>L
A:Cross-references: EMBL:Z70780; PIDN:CAA94824.1; GSPDB:GN00023; CESP:F46B6.8
A:Experimental source: clone F46B6
C:Genetics:
A:Gene: CESP:F46B6.8
A:Map position: 5
A:Introns: 35/3; 69/1; 104/1; 226/3
C:Superfamily: triacylglycerol lipase, lingual

Query Match 33.4%; Score 751; DB 2; Length 411;
Best Local Similarity 40.4%; Pred. No. 4.9e-58;
Matches 163; Conservative 72; Mismatches 138; Indels 30; Gaps 12;

QY 20 LILVAYMFORNVSVHMTKAVDPEAFNISEIIHQHGYPCPEEYEVATEDGYILSVNRIP 79
DB 10 LICTSFLLOTINA-HPDD---DPENLNNSTQIIERGWYKAEVHTVTTEDGYIILQMORIP 65
QY 80 RGLVQPKKT-----GSRPVVLLQGLVGGASNWSLNNPNSLGFILADAGFVMMGNSR 133
DB 66 YG-----KTSVTWPNGRPVILLQGLLACASDWDNLPQTQSAFVADAGFDVWLGNV 120
QY 134 GNWSRKHKTLSDIDQDFWAFSYDEMARFDLPVAVNFILQKTGOEKIYVYVSGTGMGF 193
DB 121 GTTYGRKNTKLPSETAFQFSDWMAQYDVPAMVDHVLAMTGOENLYYMGHSQCTLIME 180
QY 194 --IAFSTWPELAOKIKMYFALAPATYKHAQSPGTEKLLLPDMNMGK--LFGKKEFLYQ 248
DB 181 THLAKRDGGSFAKKIKRYFALAPAGVANKTKGFLSYFAHKFSPFDGMYELFKDFLPD 240
QY 249 TRFLRQLVILYCGOVIDL-QICSNMILLGSGFNTNNMNSRASVYAAHTLAGTSVQNILH 307
DB 241 NWITKMAKDICGASEAEKELCDNELFLIGPESDQWNSATYSSODPAGTSTQNIVH 300
QY 308 WSAVNSGELRAFDMGSETKNLEKCNQPTPVY---RVYRDMTVPMTAMWTGGQDWLSNPD 364
DB 301 WMOVRNGRVPAPDFWGGKI-NKKYGOQDTPPEYDFGAIKGTKI--HLWSDDDMLGDPD 357
QY 365 V-KMLLSEVTLNLYKHN-IPEWAHVDFTWGLDAPHRMYNEI 404
DB 358 IHDFELKELNPAVIAENVLKNEYHLDFTWGLNATFQIYDKAI 400
```

```

RESULT 8
T22675
hypothetical protein F54F3.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T22675
R:Percy, C.; Lloyd, C.
submitted to the EMBL Data Library, September 1996
A:Reference number: "19598
A:Accession: T22675
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-405<N>L
A:Cross-references: EMBL:Z79696; PIDN:CAB01973.1; GSPDB:GN00023; CESP:F54F3.3
A:Experimental source: clone F54F3
C:Genetics:
A:Gene: CESP:F54F3.3
A:Map position: 5
A:Introns: 31/3; 65/1; 274/3; 366/3
C:Superfamily: triacylglycerol lipase, lingual

Query Match 32.5%; Score 730; DB 2; Length 405;
Best Local Similarity 38.6%; Pred. No. 3.4e-56;
Matches 161; Conservative 69; Mismatches 155; Indels 32; Gaps 11;

QY 15 MEMWLLILVAYMFORNVSVHMTKAVDPEAFNISEIIHQHGYPCPEEYEVATEDGYILS 74
DB 1 MRSWSTVNLAV/LATAATVFGH---DADPEKMTTPOIIMRWGYPAMIYDVTEDGVILE 56
QY 75 VNRIPRGLVQPKKT-----GSRPVVLLQGLVGGASNWSLNNPNSLGFILADAGFDVW 128
DB 57 LHRIPYG-----KTNVTPNGKKPVVFMQHGLESCSSWVNVNLPTEAALFADAGYDVW 111
QY 129 MGNRGNWAKSKKHTLSIDQDFWAFSYDEMARFDLPVAVNFILQKTGOEKIYVYVSGQ 188
DB 112 LGNFRGNTYSMKHKLKPSHSAFMDSWDEMQQYDLPAMIEKALEVGTGDSLYYHSGQ 171
QY 189 T-TMGFIATFSTWPE---LAQKIKMYFALAPATYKHAQSPGTEKLLLPDMNMGK---LF 240
DB 172 TLTW---FSRLSEDKVGVGNKTKKFFALAPVGSVKHKGALKAFFADYFSLFEDGDFV 227
QY 241 GKKEFLYQTRFLRQLVILYCGOVIDLQICSNMILLGSGFNTNNMNSRASVYAAHTLAG 299
DB 228 GSGEFLPNMIMKLVSSEVAGLVKVEAGVCDVDFVFLIAGPESQNLNATRPVYVAHTPAG 287
QY 300 TSVONILHWSQAVNSGELRAFDMGSETKNLEKCNQPTPVYRVYRDMTVPMTAMWTGGQDWL 359
DB 288 TSTONIVHWMQVHGGVPAYDMGKT-ERGNKKHYGOANVPAYDFTTVNRPVYLYMGSDSWL 346
QY 360 SNPEDV-KMLLSEV---TNLIYKHNIPWMAHVDFTWGLDAPHRMYNEIHLMOOETN 413
DB 347 ADPTDVTDFLTLNLPSTVQNNKLDYINHLDFIWGLRAPDRIEPIIDIVRNDVLN 403
```

```

RESULT 9
H88930
protein R1LG11.14 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C:Accession: H88930
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A:Reference number: A75000; MUID:99089613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A:Accession: H88930
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-405<STO>
A:Cross-references: GB:chr.V; PIDN:AAC69088.1; PID:g2384863; GSPDB:GN00023; CESP:R1LG
A:Note: Similar to lipase; R1LG11.14
```


C:Genetics:	
A:Gene: R1LG11.14	
A:Map position: 5	
C:Superfamily: triacylglycerol lipase, lingual	
Query Match	31.5%; Score 707.5; DB 2; Length 405;
Best Local Similarity	39.0%; Pred. No. 3.2e-54;
Matches	162; Conservative 75; Mismatches 133; Indels 45; Gaps 15;
QY 19	LLLVAYMFORNVNSVHPTKAVDPEAFMNISEIIIOHGVPCPEYEVATEDGVILSVNRI 78
DB 9	LLVILA-----VHNVHAKS---DPELHMTTPOIIRWGYPMIYVTTDDGILHCHRI 59
QY 79	PRGLVOPKKT-----GSRPVLLQHLGVGGASNWSINLPNNSLGFILADAGDVWNGNS 132
DB 60	PHG-----KTNVTWPGKQPVFMQHLGCLASTDWTMLNPEQSAAFIADAGDVWLGNN 114
QY 133	RGNWRSKHKHTLSIDODEWAFSYDEMARFDLPVAVINFILQKTQGEKIYVGVSGQT-TW 191
DB 115	RGNTYSKKHKNLAKASHDFEWSDEMATYDLPAMDKVLEVTGQESLYMGHSQGTLM 174
QY 192	GFIAFTMP-----LAQIKIMYFALAPIATVKHAKSPGKTKFLLPDMIMKG---LRGKKE 244
DB 175	-----FHLSKDDGIFAKKIKKFFALAPVGSVKDIKFLGFHFFHFFLEFDGWDFVAGE 230
QY 245	FLYQTRFLRQLVIYLCGV-ILDQICSNIMLLGGFNNTNMNSRASVYAAHTLAGTSVQ 303
DB 231	FLPNWAMKLAADKICGLKIESDLCDNVCFLIAGPESDOWNSTRVPVYASHDPAGTATQ 290
QY 304	NILHWSQVNSGELRAFDMGSETKNLEKCN-QPTPVRYVRVDMT-----VPTAMWTGGQDW 358
DB 291	NIVHWIQMVRHGVPAIDMGSK-ENKKNVNFOANPEY---DFTAIGKTQIYLYWSADW 346
QY 359	LSNPEDV-KMLLSEVNLIVHKN--IPEMAHVDPIGLDAPHRYMNEIHLMOQEE 410
DB 347	LADTDTNLLRLNPAIIAQNRYFTDYNHDFVFLGRAPNDIYLPIDVICTKD 401
RESULT 10	
G89074	
protein K04A8.5 [imported] - Caenorhabditis elegans	
C:Species: Caenorhabditis elegans	
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001	
C:Accession: G89074	
R:Anonymous, The C. elegans Sequencing Consortium.	
Science 282, 2012-2018, 1998	
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology	
A:Reference number: A75000; MIMD:99069613; PMID:9851916	
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele	
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and	
A:Accession: G89074	
A:Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-411 <STO>	
A:Cross-references: GB:chr_V; PIDN:AAC48051.1; PID:g1658354; GSPDB:GN00023; CESP:K04A8.5	
C:Genetics:	
A:Gene: K04A8.5	
A:Map position: 5	
C:Superfamily: triacylglycerol lipase, lingual	
Query Match	30.8%; Score 692; DB 2; Length 411;
Best Local Similarity	36.8%; Pred. No. 7.5e-53;
Matches	153; Conservative 83; Mismatches 154; Indels 26; Gaps 10;
QY 11	VSHRMEMLLILVAYMFORNVNSVHPTKAVDPEAFMNISEIIIOHGVPCPEYEVATEDG 70
DB 5	ISDLMTWPIILLISLNSY-----SKSVDFEFLDTPELIKSWGYSYIYNTTKDG 57
QY 71	YILSVNRIPLGLVOPKKT---GSRPVLLQHLGVGGASNWSINLPNNSLGFILADAGDV 127
DB 58	FIELEHRIPIYGREVPTSSDVNNRPVIFLQHGFLCSFSDWVNSPQSGAFVADAGDV 117
QY 128	WMGNSRGNSRKHKHTLSIDQD-EFWAFSYDEMAREFDLPVAVINFILQKTQGEKIYVGV 186

Db 118	WLGNFRGNTYSRKHVSLNPKDKPFWDSWDQISEYDLPAMIKALISGOESLYTGF 177
QY 187	QGSTMGFIAPSTMPPELAQKIKMYFALAPIATVKHAKSPGKTKFLLPDMIMKG---F 240
DB 178	LGSLTWPAKLSTDPKFSRKIKKYFALAPIGSIKHAHG---VFELGRHFGKDYBEYVKKH 234
QY 241	GKKEFLYQTRFLRQLVIYLCGV-ILDQICSNIMLLGGFNNTNMNSRASVYAAHTLAG 299
DB 235	GSDELFGSSLLFKKIVKYTCGLDTEEFCSDDITLLFIGNANENWQTRIPVYLAHTPAG 294
QY 300	TSVQNILHWSQVNSGELRAFDMGSETKNLEKCNQPTPVRYVRVDMT-VPTAMWTGGQDW 358
DB 295	SSSNVMAHLDMQFSYGVPTDFMGEE-KNLKAYGQKLPQYVNFETGIADVPYLFWSDDW 353
QY 359	LSNPEDV-KMLLSEVNLIVHKN--IPEMAHVDPIGLDAPHRYMNEIHLMOQEE 411
DB 354	LSYKQDLLEWFLPAQLNSQVVOGSRFRIENYHLEHPIWGTNNAASQVINYITGILQDD 409
RESULT 11	
T20480	
hypothetical protein F01G10.7 - Caenorhabditis elegans	
C:Species: Caenorhabditis elegans	
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000	
C:Accession: T20480	
R:Hembry, C.	
submitted to the EMBL Data Library, October 1996	
A:Reference number: Z19280	
A:Accession: T20480	
A:Status: preliminary; translated from GB/EMBL/DBJ	
A:Molecule type: DNA	
A:Residues: 1-426 <WIL>	
A:Cross-references: EMBL:Z81055; PIDN:CAB02896.1; GSPDB:GN00022; CESP:F01G10.7	
A:Experimental source: Clone F01G10	
C:Genetics:	
A:Gene: CESP:F01G10.7	
A:Map position: 4	
A:Introns: 29/3; 64/1; 186/3; 347/2	
C:Superfamily: triacylglycerol lipase, lingual	
Query Match	29.2%; Score 655.5; DB 2; Length 426;
Best Local Similarity	35.9%; Pred. No. 1.3e-49;
Matches	140; Conservative 85; Mismatches 144; Indels 21; Gaps 9;
QY 20	LILVAYMFORNVNSVHPTKAVDPEAFMNISEIIIOHGVPCPEYEVATEDGVILSVNRI 79
DB 6	LFULILSFSINL-----SLAIDDECYMTVPEIGKHFGYSEVHLVRTTDEYILELHRIP 59
QY 80	RGVLQVPKKTGSRPVLLQHLGVGGASNWSINLPNNSLGFILADAGDVWNGNSRGNAWS 138
DB 60	CKQNEKCDRSKRPRIVFMQHLGLADGFSWIPNLANSAGFVFADAGDFIWANSRGTPAS 119
QY 139	RKHKHTSIDODEWAFSYDEMAREFDLPVAVINFILQKTQGEKIYVGVSGQTMGFIAPST 198
DB 120	QKHIGYGPENQKFNWTFWQMSDFDLTASVDLVLETKQEFLLYLLGHSGQTMWINSRLAE 179
QY 199	MPELAQIKMYFALAPIATVKHAKSP-----GTFKLLPDMIMKGILFGKKEP--LYQTRFL 252
DB 180	NPEFAKKIRHFHALAPVATVSHIGLFGLEGTKEFTYAEI---LLGLRPLSPISIPRTV 235
QY 253	RLQVLYLCGVILDQICSNIMLLGGFNNTNMNSRASVYAAHTLAGTSVQNILHWSQAV 312
DB 236	QKMSISYMSRFFQNICTLDIGFDG-NEKMFNSRVGVYLCHTPATSVKDLQHNIOLV 294
QY 313	NSGELRAFDMGSETKNLEKCNQPTPVRYVRVDMTVPVYVYVYVYVYVYVYVYVYVYVY 371
DB 295	KSQTVSNFXYGTD-GNIIEXQPTPEYDVTQINTPTVLYWSRDLADTQDIRDSILSK 353
QY 372	VTNLIYHK-NIPEWAHVDFIWLGDAPHRYM 400
DB 354	MNKTAGSLPELPHYSMDHDFVFGTHAAFDLY 383

RESULT 12

eggs-specific protein - silkworm
C:Species: Bombyx mori (silkworm)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 26-Feb-1998
C:Accession: JT0949; A28527
R:Sato, Y.; Yamashita, O.
submitted to JIPID, September 1991
A:Reference number: JT0949
A:Accession: JT0949
A:Molecule type: DNA
A:Residues: 1-559 <SAT>
A:Experimental source: larva
R>Note: this protein is a homotrimer
R:Indrasith, L.S.; Sasaki, T.; Yamashita, O.
J. Biol. Chem. 263, 1045-1051, 1988
A:Title: A unique protease responsible for selective degradation of a yolk protein in *Bombix mori*
A:Reference number: A28527; PMID:8087166; PMID:3275655
A:Accession: A28527
A:Molecule type: protein
A:Residues: 19-40; 133-144, 'T', 146-152; 229-248 <IND>
C:Genetics:
A:Map position: 19
C:Keywords: egg yolk; homotrimer
F:132-133/Cleavage site: Lys-Asn (egg 30.5K cysteine proteinase) #status experimental
F:228-229/Cleavage site: Arg-Asp (egg 30.5K cysteine proteinase) #status experimental

Query Match 25.6%; Score 574.5; DB 2; Length 559;
Best Local Similarity 33.2%; Pred. No. 2.5e-42;
Matches 132; Conservative 82; Mismatches 152; Indels 31; Gaps 9;

QY 32 NSVH--MPTKAVDPEAF-MNISIIHQGVPCIEYEVATEDGYLSVNRIPRGLVQPKK 87
DB 171 NTFDAISETORENNEDFHLNATELLKKHQYVVEHVATDDGGYHLTVLRIP-PTHQTRD 229

QY 88 TGSRPVLLQGLVGGASNWSINLNNLSLGLADAGFDVGMNWSRGNWSRKHKLTSID 147
DB 230 DKKKPVALLHGLLGSADDMLLMGPSKSLAYMLSDAGYDVWLVGNVGRNKSRSVSKRHPA 289

QY 148 QDEFWASDEMAFRDLPVAVINFLQGTQKQIYVYVYSGQTTMGFTAFSTWPELAQKIK 207
DB 290 LNDPWFKEFNDEIALHDLPAITDHYDLSGQERLHYIGHSQGATTFALMSQPSYNEKIV 349

QY 208 MYPALAPLATVKKHAKS-----PGTKFL-LLPDMTKGLFGKKEFLYQTRFLQLVIYL 259
DB 350 SHWALSPIVYMYVRSPLFRMIATSFYQYIHDQVGHGAPEPKHLIET-----F 400

QY 260 CGQVILDQ-----ICSNIMLLGGFTNNMNSRASVYAAHTLAGTSVQNTLHWSQAVNS 314
DB 401 GGAACREKLGCRHVCNNLNVVISGVNYNQDADIVPVVMAHLPGATTSARVMKQYGNVAS 460

QY 315 GELRAFQWSETKLEKCNQPTPVRYVRDVTVPMTAMTGGQDMLSPEDVKMLLSEVTN 374
DB 461 HFRKYNIGAEI-NMKYVYGAEPSPSYDLSKVSAPVNLVSHSDAWLAHPKQVKEKLQENLPN 519

QY 375 LIYKNTPEWNAH---VDIFWGLDAPHRMYNEIILMQ 408
DB 520 VKQSFEVPEQOHFTDLDFQFSKKAPDVTYQKLMMNQ 556

RESULT 13
T31611
triglyceride lipase-cholesterol esterase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: T31611
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.
submitted to the EMBL Data Library, February 1998
A:Reference number: 221862
A:Accession: T31611
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-443 <WOO>
A:Cross-references: EMBL:AL021748; PIDN:CAA16863.1; GSPDB:GN00067; SPDB:SPBC16A3.12c
A:Experimental source: strain 972h-; cosmid c16A3
C:Genetics:
A:Gene: SPDB:SPBC16A3.12c
A:Map position: 2
A:Introns: 11/2; 152/3; 282/1; 398/3
C:Superfamily: triacylglycerol lipase, lingual

Query Match 22.5%; Score 505; DB 2; Length 443;
Best Local Similarity 29.8%; Pred. No. 2.4e-36;
Matches 125; Conservative 91; Mismatches 154; Indels 52; Gaps 13;

QY 9 WIVSHRMEM--LLILVAYMFQRNVNSVHMPTRKAVDPEAFWNISEIIHQGYCEEYEA 66
DB 35 YMIPERVREVRRIIVLYHYIYSSKTTDGMTDAV--QKCRNIYEICEAFGRVVEHLVR 92

QY 67 TEDGYILSVNRIPRGLVQPKKT-GSRPVLLQGLVGGASNWSI-NLPNNSLGLADAG 124
DB 93 TDQNFILCHRI-----THPKSQSHKREVYCHHGLMTNSELWAVAVNESRSLPFLTESG 148

QY 125 FDVWNGNSRGNWSRKHKLTSIDODEFWAFSYDEMAFRDLPVAVINFLQGTQKQIYVYG 184
DB 149 YDWLGNNGNKYSRKHITYPKDEEFNFSLDMMAFDIPDVTVDYLLRETGREKLNYIG 208

QY 185 YSQGTTMGFAFSTMPPELAQKIKRYFALAPATYVKKHAKSPGTRFLL--PDMIKGLFGK 242
DB 209 FSGTAQAMAALSINPLNDKVNIFIGLAPAYAPKFSNVFYVDYIVKVNPKIMYH-LFGR 267

QY 243 KEFLYQTRFRLQVLYICGGVLDQICSNIMLLGGFTNNMNSRASVYAAHTLAGTSV 302
DB 268 RCLLPSVTFWJN---ICYPPIFKIVDSLKILFNLDLSNISLNQKLCGYAHLYSFSV 323

QY 303 QNIIHWSQAVNSGELRAFD-----WGSE-----TKNLEKCNQPTPVRYVRDMT 346
DB 324 KSVVHLLQIINKCTFQLYDDMALLAGYGRHYQVLPFTNNI-KC----- 368

QY 347 VPTAMWTGGQDWLSPEDVKMLLSEVTNLIYKNIPEWNAHVDFTWGLDAPHRMYNEIHL 406
DB 369 -PMILLGGRTDLINMEVMTALPPHAKV---SIAHYEHLDFLWGQDVKEEVPVVIDA 424

QY 407 MQ 408
DB 425 LK 426

RESULT 14

T31611
hypothetical protein Y50E8A.g - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T31611
R:Steward, C.
submitted to the EMBL Data Library, September 1999
A:Reference number: 221047
A:Accession: T31611
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1585 <WIL>
A:Cross-references: EMBL:AL117200; NID:el549770; PIDN:CAB55050.1; CESP:Y50E8A.g
A:Experimental source: clone Y50E8A
C:Genetics:
A:Gene: CESP:Y50E8A.g
A:Introns: 25/3; 60/1; 133/2; 217/3; 270/3; 337/2; 400/1; 746/2

Query Match 21.3%; Score 478.5; DB 2; Length 1585;
Best Local Similarity 30.6%; Pred. No. 3.1e-33;
Matches 118; Conservative 67; Mismatches 172; Indels 29; Gaps 9;

QY 44 EAFWNISEIIHQGYPCIEYEVATEDGYLSVNRIPRGLVQPKKTG--SRPVLLQGLV 101
DB 20 EETLDAADTISHYGYTEKHVYTTDDGYTQVQLIPVGRDRDRSLGCKSRKRPVFFMHGLF 79

Search completed: February 3, 2003, 13:50:26
Job time : 23 secs

Qy	39	KAVDPFAFMNISEIIHQGVPCBEYEVATEDGYILSVNRIPR-----GLVQPKTKGSRP	92
Db	73	KAIHDA--DIREMKISGIVYVEDHLVRTEDDYILCIHRISKSDSPGRIGSPHPKLL--P	127
Qy	93	VVLLQHLGVGCASNWISNL-PNNSLGFIADAGDFVWNGSRNGNAWSRKHKTLSIDQDFE	151
Db	128	VYVCHHGLLNSEVWVCNVDPRNCLYDFLVNKGIDYVWLGNRGNKYSRHOLFDFSDTKEF	187
Qy	152	WAFSYDEMARFDLPVINFILQKTGOEKYYVYVGSOGTTMGITASTMPELAQKIKMFA	211
Db	188	WDFSIDFAQYIDPTDYLITKSGOTKLTGYFGSGTAQAFASLSIHPLLNKDKINSLSIA	247
Qy	212	LAPIATVVKHAKSPGTFLLLPDMNWK-----LFGKKEFYLQTRFLROLVYLCGQVI	264
Db	248	LAP-----AISPGLHNRVVDAFVKARPSILFELEGRKSILPSAGFWQSFL-----APXF	297
Qy	265	LDOICSNMILLGGFTNNMNSRASVYAAHTLAGTSYQNIHLWSQAVNSGELRAFD---	321
Db	298	FDVAVLAYCLSQLFNWMSQONISSYQRLVSFAHLYSYTTSVKCLVHWFIQMSAEFRWYDNDQ	357
Qy	322	WGSE--TKNLEKCNQTPYRYVRDVMVTPTAMWTGGQDWLSPEDVKMLLSEVTLNIIYHK	379
Db	358	LGHDFELKYYKAAKFTP-----NNRTPIYLWGGSDSL---VDIOAMLNALPAEVEHV	408

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:44:54 ; Search time 14 Seconds
(without alignments)
1253.178 Million cell updates/sec

Title: US-09-333-159-47

Perfect score: 2247

Sequence: 1 MLETLRQWIVSHRMEMWLL.....IHLMQQETNLSQGRCAVL 423

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1289	57.4	399	1 LICH_HUMAN	P38571 homo sapien
2	1220.5	54.3	397	1 LICH_RAT	Q64194 rattus norv
3	1197.5	53.3	397	1 LICH_MOUSE	Q920m5 mus musculu
4	1164.5	51.8	395	1 LIPG_MOUSE	P04634 rattus norv
5	1154	51.4	398	1 LIPG_CANFA	P80035 canis fami
6	1139	50.7	397	1 LIPG_BOVIN	Q29458 bos tauris
7	1139	50.7	397	1 LIPG_HUMAN	P07098 homo sapien
8	663	29.5	394	1 LIP3_DROME	O46108 drosophila
9	546	24.3	439	1 LIP1_DROME	O46107 drosophila
10	422.5	18.8	548	1 TGL1_YEAST	P34163 saccharomyc
11	126	5.6	987	1 YD94_METJA	Q58789 methanococc
12	105	4.7	430	1 ORC2_CAEEL	Q21037 caenorhabd
13	101	4.5	436	1 PAPA_CAVPO	P70683 c platelet-
14	96.5	4.3	287	1 Y193_HAEIN	Q57427 haemophilus
15	95.5	4.3	1259	1 L1N1_HUMAN	P08547 homo sapien
16	93.5	4.2	278	1 PRXC_STRAU	O31168 streptomyc
17	93	4.1	569	1 PYRD_PLAFA	Q08210 plasmodium
18	92	4.1	748	1 PTIP_ECOLI	P37177 escherichia
19	92	4.1	748	1 PTIP_SALTY	P37178 salmonella
20	91	4.0	370	1 P2C2_SCHPO	Q09172 schizosacch
21	90.5	4.0	1037	1 ACRD_ECOLI	P24172 escherichia
22	90	4.0	277	1 BPA2_STRAU	P29715 streptomyc
23	90	4.0	1122	1 YG3C_YEAST	P53280 saccharomyc
24	90	4.0	1136	1 POLG_OMV	P20234 o genome po
25	90	4.0	2083	1 DYSF_MOUSE	Q9esd7 mus musculu
26	89	4.0	330	1 EXOA_RHIME	P33691 rhizobium m
27	88.5	3.9	185	1 YCZH_BACSU	O31482 bacillus su
28	88.5	3.9	869	1 CFAC_ECOLI	P25733 escherichia
29	87	3.9	473	1 SCRB_LACLA	Q04937 lactococcus
30	86.5	3.8	382	1 YDHF_HAEIN	P44861 haemophilus
31	86.5	3.8	685	1 INVA_SALTY	P35657 salmonella
32	86.5	3.8	1882	1 Y468_MYCPN	P75109 mycoplasma
33	86	3.8	525	1 ACU8_NEUCR	P15937 neurospora

34	86	3.8	1866	1 VGNB_CPMV	P03500 cowpea mosa
35	85.5	3.8	428	1 YE63_SCHPO	O14249 schizosacch
36	85	3.8	429	1 ARE2_ECOLI	P52146 escherichia
37	85	3.8	837	1 UBPA_DICDI	P54201 dictyosteli
38	85	3.8	4128	1 PRKD_MOUSE	P97313 mus musculu
39	84.5	3.8	875	1 AMD2_XENLA	P12890 xenopus lae
40	84	3.7	451	1 YB27_YEAST	P38295 saccharomyc
41	84	3.7	804	1 SYL_BACSU	P36430 bacillus su
42	83.5	3.7	350	1 EL3B_PRUPE	P52408 prunus pers
43	83.5	3.7	375	1 NDR3_MOUSE	Q9qvf9 mus musculu
44	83.5	3.7	389	1 AAT_PYRAB	Q9v012 pyrococcus
45	83.5	3.7	456	1 YP95_YEAST	Q02891 saccharomyc

ALIGNMENTS

RESULT 1
LICH_HUMAN
ID LICH_HUMAN STANDARD: PRT: 399 AA.
AC P38571; Q16529; Q96EJ0;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lysosomal acid lipase/cholesteryl ester hydrolase precursor
DE (EC 3.1.1.13) (LAL) (Acid cholesteryl ester hydrolase) (Sterol
DE esterase) (Lipase A) (Cholesteryl esterase).
GN LIPA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 196-212; 277-297 AND 305-315.
RX MEDLINE=92042192; PubMed=1718995;
RA Anderson R.A., Sando G.N.;
RT "Cloning and expression of cDNA encoding human lysosomal acid
RT lipase/cholesteryl ester hydrolase. Similarities to gastric and
RT lingual lipases";
RL J. Biol. Chem. 266:22479-22484(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94155897; PubMed=8112342;
RA Ameis D., Merkel M., Eckerskorn C., Greten H.;
RT "Purification, characterization and molecular cloning of human
RT hepatic lysosomal acid lipase";
RL Eur. J. Biochem. 219:905-914(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=96363957; PubMed=8725147;
RA Du H., Witte D.P., Grabowski G.A.;
RT "Tissue and cellular specific expression of murine lysosomal acid
RT lipase mRNA and protein";
RL J. Lipid Res. 37:937-949(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP VARIANT CSD/WD PRO-200.
RX MEDLINE=94195814; PubMed=8146180;
RA Anderson R.A., Byrum R.S., Coates P.M., Sando G.N.;
RT "Mutations at the lysosomal acid cholesteryl ester hydrolase gene
RT locus in Wolman disease";
RL Proc. Natl. Acad. Sci. U.S.A. 91:2718-2722(1994).
RN [6]
RP VARIANT3 CSD ARG-129 AND PRO-129.
RX MEDLINE=98295576; PubMed=9633819;
RA Ries S., Buechler C., Schindler G., Aslanidis C., Ameis D., Gasche C.,
RA Jung N., Schambach A., Fehrer P., Vanier M.T., Belli D.C.,

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: S81497; AAB36043.2; -
 DR InterPro: IPR000733; Abhydrolase.
 DR InterPro: IPR000734; Lipase.
 DR InterPro: IPR000379; Ser_estrs_site.
 DR Pfam: PF00561; abhydrolase; 1.
 DR PROSITE: PS00120; LIPASE_SER; FALSE_NEG.
 KW Hydrolase; Lipid degradation; Glycoprotein; Signal; Lysosome.
 FT SIGNAL 1 19
 FT CHAIN 20 397
 FT
 FT ACT_SITE 172 397
 FT ACT_SITE 372 372
 FT CARBOHYD 34 34
 FT CARBOHYD 99 99
 FT CARBOHYD 159 159
 FT CARBOHYD 271 271
 FT CARBOHYD 319 319
 SQ SEQUENCE 397 AA; 45186 MW; 97A38595A0523947 CRC64;
 Query Match 54.3%; Score 1220.5; DB 1; Length 397;
 Best Local Similarity 58.2%; Pred. No. 2.5e-98;
 Matches 220; Conservative 67; Mismatches 84; Indels 7; Gaps 3;
 QY 37 PT---RAVDPEAFPMNISEIIHQGYPCPEYEVATEDGYILSVNRIPRGLVQPKKTSRVP 93
 DB 20 PTCISAVDPEANMNVTEIIMHWGP--EHSVQTGDIYLGVRIPHGRKNQFDKGPVP 77
 QY 94 VLQ--HGLVGGASNISLPPNNSLGFILADAGFDVWNGNSRGNWSRKHKLTSIDODEF 151
 DB 78 VYLQWRHGFGLADSSNVTNIDNLSGLFILADAGFDVWNGNSRGNWSRKHKLTSVSDOEY 137
 QY 152 WAFSYDEMAREDPVAVINFLQKTOEKIYVYVSGTGMFTAFSTMPELAKIKMYFA 211
 DB 138 WAFSDEMAREDPVAVINFLQKTOEKIYVYVSGTGMFTAFSTMPELAKIKMYFA 197
 QY 212 LAPAVTKHAKSPGTFLLLPDMNIMKGLFGKKEFLYQTRFLRQVILYLCQVILQICSN 271
 DB 198 LAPVLSINFASGPMVGLRPLDLLLEDLFGKQFLQPSAMVKNLSTHICTHVMKELCAN 257
 QY 272 IMLLGGFNPNMNSRASVYAHTLAGTSVQNLHWSQAVNSGELRAPDWSGETNLEK 331
 DB 258 IFFLICGFNEKLNMSRDVYTHCPAGTSVQNVHWTQVYKHLQAFDWSGSDKNYFH 317
 QY 332 CNOPTPVRYRVRDVTPTAMWTGQDWSNPEDVKMLSEVTNLVYKKNIPKAWHVDFIW 391
 DB 318 YNOSYPLYSIKDMQPLTALMSGCKDLADTSDINILLRTEIPLTVTHKNIPEDWHDIFIW 377
 QY 392 GLDAPHRMYNEIHLMOQ 409
 DB 378 GLDAPWRLYNEVYSLMKK 395

RESULT 3

LICH_MOUSE STANDARD; PRT; 397 AA.
 AC 0920M5;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Lysosomal acid lipase/cholesterol ester hydrolase precursor
 DE (EC 3.1.1.13) (LAL) (Acid cholesterol ester hydrolase) (Sterol
 DE esterase) (Lipase A) (Cholesterol esterase).
 GN LIPA OR LIPI.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-C57BL/6 X CBA; TISSUE=Liver;
 RX MEDLINE=96363957; PubMed=8725147;
 RA Du H., Witte D.P., Grabowski G.A.;
 RT "Tissue and cellular specific expression of murine lysosomal acid
 RT lipase mRNA and protein";
 RL J. Lipid Res. 37:937-949(1996).
 CC -!- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL
 CC ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-
 CC MEDIATED ENDOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN
 CC MEDIATING THE EFFECT OF LDL (LOW DENSITY LIPOPROTEIN) UPTAKE ON
 CC SUPPRESSION OF HYDROXYMETHYLGUTARYL-COA REDUCTASE AND ACTIVATION
 CC OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION (BY
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: A sterol ester + H(2)O -> a sterol + a fatty
 CC acid.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- TISSUE SPECIFICITY: Expressed at low levels in most tissues. High
 CC level expression is found in hepatocytes and splenic and thymic
 CC cells. Very high level expression is observed in cells of the
 CC small intestinal villi, the zona fasciculata and tubularis of
 CC the adrenal cortex, pancreatic acini, and renal tubular
 CC epithelium.
 CC -!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: 231689; CAAB3494.1; -
 DR MGD; MGI:96789; Lip1.
 DR InterPro: IPR000073; Abhydrolase.
 DR InterPro: IPR000734; Lipase.
 DR InterPro: IPR000379; Ser_estrs_site.
 DR Pfam: PF00561; abhydrolase; 1.
 DR PROSITE: PS00120; LIPASE_SER; 1.
 KW Hydrolase; Lipid degradation; Glycoprotein; Signal; Lysosome.
 FT SIGNAL 1 19
 FT CHAIN 20 397
 FT
 FT ACT_SITE 172 397
 FT ACT_SITE 372 372
 FT CARBOHYD 34 34
 FT CARBOHYD 99 99
 FT CARBOHYD 159 159
 FT CARBOHYD 271 271
 FT CARBOHYD 319 319
 SQ SEQUENCE 397 AA; 45551 MW; F886C39E1CCFA91F CRC64;
 Query Match 53.3%; Score 1197.5; DB 1; Length 397;
 Best Local Similarity 55.6%; Pred. No. 2.5e-96;
 Matches 217; Conservative 70; Mismatches 100; Indels 3; Gaps 1;
 QY 23 VAYNFORNVNSVHMPT---KAVDPEAFPMNISEIIHQGYPCPEYEVATEDGYILSVNRIP 79
 DB 6 LVEVFTIGILSRVPTGTVSAVDPEVNMMVTETIIMWGYPCGEHSLVLTGDIYLSHRIP 65
 QY 80 RGLVQPKKTSRVPVVLQHLVGGASNISLPPNNSLGFILADAGFDVWNGNSRGNWSR 139
 DB 66 RGNHFGKGRPVVYLYQHLGLADSSNVTNIDNLSGLFILADAGFDVWNGNSRGNWSL 125
 QY 140 KEKTLSTDDQEFWAFSYDEMAREDPVAVINFLQKTOEKIYVYVSGTGMFTAFSTM 199
 DB 126 KEKTLSTVSDQEFWAFSDEMAREDPVAVINFLQKTOEKIYVYVSGTGMFTAFSTM 185
 QY 200 PELAKIKMYFALAPIATVAKHAKSPGTFLLLPDMNIMKGLFGKKEFLYQTRFLRQVILY 259
 DB 186 PELAKIKMYFALAPIATVAKHAKSPGTFLLLPDMNIMKGLFGKKEFLYQTRFLRQVILY 245
 QY 260 CQVILQICSNIMLLGGFNPNMNSRASVYAHTLAGTSVQNLHWSQAVNSGELRA 319

```

Db 246 CTHVIMKELCANVFLLCGNEKLNLSRVDVYTHCPAELLVQMLHWGVQYKRLQA 305
QY 320 FDMGSEKTKLEKCNQPTPVRYVRVDMVPTAMMTGGODWLSNPEDVKMLLSEVTNLVHK 379
Db 306 FDMGSSKNTFYHNSPPSYNKNMRLPTALWGGGRDLADINDITILLTQIKPLVYHK 365
QY 380 NIPWAHVDFIWLGDAPHRMYNEIHLMOQ 409
Db 366 NIPEDWHLDFIWLGDAPKLYDEIISLMKK 395

RESULT 4
LIPG_RAT
ID LIPG_RAT STANDARD; PRT; 395 AA.
AC P04634;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Triacylglycerol lipase, lingual precursor (EC 3.1.1.3) (Lingual
lipase).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85215587; PubMed=3839077;
RA Docherty A.J.P., Bodmer M.W., Angal S., Verger R., Riviere C.,
RA Lowe P.A., Lyons A., Entage J.S., Harris T.J.R.;
RT "Molecular cloning and nucleotide sequence of rat lingual lipase
cDNA."
RL Nucleic Acids Res. 13:1891-1903(1985).
CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O -> diacylglycerol + a
fatty acid anion.
CC -1- MISCELLANEOUS: LINGUAL LIPASE IS SECRETED BY THE SEROUS (VON
EBNER'S) GLANDS AT THE BACK OF THE RAT TONGUE.
CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02309; CAA26179.1; -
DR EMBL; A01157; CAA00136.1; -
DR PIR; A23045; LIRTT.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00561; abhydrolase.1.
DR PROSITE; PS00120; LIPASE_SER; 1.
KW Hydrolase; Lipid degradation; Glycoprotein; Signal.
FT SIGNAL 1
FT CHAIN 18
FT ACT_SITE 19 395
FT ACT_SITE 171 171
FT ACT_SITE 342 342
FT ACT_SITE 371 371
FT DISULFID 245 254
FT CARBOHYD 33 33
FT CARBOHYD 68 68
FT CARBOHYD 98 98
FT CARBOHYD 270 270
SQ SEQUENCE 395 AA; 44588 MW; E501854A923522EA CRC64;

Query Match 51.8%; Score 1164.5; DB 1; Length 395;
Best Local Similarity 54.9%; Pred. No. 1.8e-93;
Matches 21; Conservative
QY 17 MWLLILVAYMFO-RNVNSVHMPTKAVDPAPMNISETIIHQGPCPEYEVEVATEDGYILSV 75

```

```

Db 1 MWLLITSVISTEGGAHGLFGKLGPNANNNISQMITYWGYPQCEYEVVTEGYILGV 60
QY 76 NRIRPGLVQPKTKSRPVVLLQHLGVGASWISNLPNNSLPGFLADAGFDVWNGNSRGN 135
Db 61 YRIRPHGKNNSENIGKRPVYVYLOHGLIASATNWIANLPNNSLAFMLADAGYDVWILGNSRGN 120
QY 136 AWSRKHKTLSIDODEFWAFSDEMARFDPVAVINFILQKTGOEKIYVYVYSGQTTMGFIA 195
Db 121 TWSRKNVYSPDSVEFWAFSDEMAKYDLPATINFIQKTGOEKIHYVGHSGQTTMGFIA 180
QY 196 FSTMPELAQIKMYFALAPIATVKHAKSPGTFLLLPDMIKGLFGKKEFLYOTFRRLQOL 255
Db 181 FSTNPTLAKKIKTEVALAPVATVYQSPKKISFIPTFLEKLMFGKMKMELPHTYFDDFL 240
QY 256 VIYLCGGVILDQICSNIMLLGGTNNMMSRASYVAAHTLAGTSVQNIHLHWSQVNSG 315
Db 241 GTEVCSREVLDLCSNTLFIQFGDKNLNYSFVYLGHNPAQTSVQDFLHWAQLVRSG 300
QY 316 ELRAFDMGSETKLEKCNQPTPVRYVRVDMVPTAMMTGGODWLSNPEDVKMLLSEVTNL 375
Db 301 KFOAFNMGWSPQNMHLHYNQKTPPEYDVSAMTPVYVWNGNDILADPODVAMLLPKLSNL 360
QY 376 IYHKNIPWAHVDFIWLGDAPHRMYNEIHLMOQ 410
Db 361 LFHKEILAYNHLDFIWMADAPQEVYNEMISMAED 395

RESULT 5
LIPG_CANFA
ID LIPG_CANFA STANDARD; PRT; 398 AA.
AC P80035; O02857;
DT 01-NOV-1991 (Rel. 20, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Triacylglycerol lipase, gastric precursor (EC 3.1.1.3) (Gastric
lipase) (GL).
GN LIPF.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Stomach;
RA Vaganay S., Jolliff G., Bertaux O., Toselli E., Devignes M.D.,
RA Benicourt C.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 20-59.
RX MEDLINE=92037652; PubMed=1935982;
RA Carriere F., Moreau H., Raphael V., Laugier R., Benicourt C.,
RA Junien J.-L., Verger R.;
RT "Purification and biochemical characterization of dog gastric
lipase."
RL Eur. J. Biochem. 202:75-83(1991).
CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O -> diacylglycerol + a
fatty acid anion.
CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y13899; CAA74198.1; -
DR PIR; S19539; S19539.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.

```

DR Pfam: PF00561; abhydrolase; 1.
 DR PROSITE; PS00120; LIPASE_SER; 1.
 KW Hydrolase; Lipid degradation; Glycoprotein; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 398 TRIACYLGLYCEROL LIPASE, GASTRIC.
 FT ACT_SITE 172 172 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 372 372 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 39 39 I -> T (IN REF. 2).
 SQ SEQUENCE 398 AA; 45130 MW; E04D62F7518E386C CRC64;

Query Match 51.4%; Score 1154; DB 1; Length 398;
 Best Local Similarity 54.3%; Pred. No. 1.5e-92;
 Matches 216; Conservative 71; Mismatches 99; Indels 12; Gaps 5;

Qy 17 MWLLILVAYMFORNVNHH-----MPTKAVDPFAFMNISEIIQHOGYPCPEEYEVATEDG 70
 Db 1 MWLLTAAASVIS-TLGTTHGLFGKLHPT---NPEVTNNISQIMITYGYPABEYEVATEDG 56
 Qy 71 YILSVNRIPRLVQPKTKGSRPVLLQHLGVGASNWSLNPNNSLGFILADAGFDVWVG 130
 Db 57 YILGIDRIPYGRKNSENTGRPRVAFQLHGLASATNWSLNPNNSLAFILADAGFDVWVG 116
 Qy 131 NSRGNAWRKHTLSIDODEFWAFSYDEMARDLPVAVINFLQKTQGEKIYYVGSQGT 190
 Db 117 NSRGNTWARRNLVSPDSVEFWAFSDEMAYDLPATIDFILKTKGQDKLHVHGSQGT 176
 Qy 191 MGTAFTSMPELAQIKMYFALAPATVATVKKHAKSPGTEKLLPDMMIKGLFKKEFLYQTR 250
 Db 177 IGFATFSTNPKLAKRIKTEYALAPVATVYKTTETLNLKMLVPSFLFKIFGNKIP-YPHH 235
 Qy 251 FLRO-LVTVLGVQVILDOICSNIMLLGGFNTNMNMSRASVYAAHTLAGTSVQNLHWS 309
 Db 236 FDDQLATEVCSRETVDLCSNALFICGFDMLNLSRLDVLVSHNPAGTSVQNLHWS 295
 Qy 310 QAVNSGELARFDWGSSEKTKNLEKCNQPTPVRYVRVDMTVPYTAAMTGGQDNLSPEDVKMLL 369
 Db 296 QAVKSGKFOAFDWGSPVQNMHHYQSMPPYNTLDMHVPYIAVWNGNDLLADPHDVL 355
 Qy 370 SEVTNLIYKHNIPEAHVDFIWLADAPHRMNEIILHM 407
 Db 356 SKLPNLIYHRKIPPYNHLDFIWAMDAPQAVYNEIVSM 393

RESULT 6
 LIPG_BOVIN STANDARD; PRT; 397 AA.
 ID LIPG_BOVIN
 AC Q29458;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Triacylglycerol lipase, pregastric precursor (EC 3.1.1.3) (Pregastric
 DE lipase) (GL) (Pregastric esterase) (PGE).
 GN LIPF.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Tongue serous gland;
 RX MEDLINE=95011625; PubMed=7926811;
 RA Timmermans M.Y.J., Kupers L.P., Teuchy H.;
 FT "The cDNA sequence encoding bovine pregastric esterase.";
 RL Gene 147:259-262(1994).
 CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O -> diacylglycerol + a
 CC fatty acid anion.
 CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; L26319; AAA57037.1; -
 DR InterPro: IPR000073; Abhydrolase.
 DR InterPro: IPR000734; Lipase.
 DR InterPro: IPR000379; Ser_estrs_site.
 DR Pfam: PF00561; abhydrolase; 1.
 DR PROSITE; PS00120; LIPASE_SER; 1.
 KW Hydrolase; Lipid degradation; Glycoprotein; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 397 TRIACYLGLYCEROL LIPASE, PREGASTRIC.
 FT ACT_SITE 171 171 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 371 371 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 326 326 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 397 AA; 45231 MW; F68977DED585EE36 CRC64;

Query Match 50.7%; Score 1139; DB 1; Length 397;
 Best Local Similarity 53.7%; Pred. No. 3e-91;
 Matches 212; Conservative 68; Mismatches 113; Indels 2; Gaps 2;

Qy 18 WLLILVAYMFORNVNHHMPTKAVDPFAFMNISEIIQHOGYPCPEEYEVATEDGYILSVNR 77
 Db 3 WLLVTVCFIHMSGNAFCFLGKIAKNPEASNMVYSQIMSYGYPSEMHKVITADGYILOVYR 62
 Qy 78 IPRGLVQPKTKGSRPVLLQHLGVGASNWSLNPNNSLGFILADAGFDVWNGSRGNW 137
 Db 63 IPHGKNANHLGQRPVFLQHLGSLGATNWSLNPNNSLGFILADAGFDVWNGSRGNW 122
 Qy 138 SRKHTKLSIDODEFWAFSYDEMARDLPVAVINFLQKTQGEKIYYVGSQGTMGTFIAT 197
 Db 123 AQEHLVYSPDSPEFWAFSDEMAYDLPSTIDFILRTQKQLHYVHSGTGTIGTFIAT 182
 Qy 198 TPELAQIKMYFALAPATVATVKKHAKSPGTEKLLPDMMIKGLFKKEFLYQTRFLQ-LV 256
 Db 183 TSPTLAELKIKVEYALAPVATVYKTSFLFNKLALPHFLFKIIFGDKMF-YPHTFLOFLG 241
 Qy 257 IYLCQVILDOICSNIMLLGGFNTNMNMSRASVYAAHTLAGTSVQNLHWSQAVNSGE 316
 Db 242 VEMCSRETLDVLCCKNALFAITGVNDKNFNNSRLDVYIAHNPAGTSVQNLHWSQAVNSG 301
 Qy 317 LRAFDMGSETKNLEKCNQPTPVRYVRVDMTVPYTAAMTGGQDNLSPEDVKMLLSEVTNLI 376
 Db 302 FOAFDWGAPYQNLHMYHQPPIYNTLTMNVPIAVNSADNLLADPDQDVFLLSKLSNLI 361
 Qy 377 YHKNIPWAHVDFIWLADAPHRMNEIILMQOEE 411
 Db 362 YHKEIPYNHLDFIWAMDAPQEVYNEIVSLMAEDK 396

RESULT 7
 LIPG_HUMAN STANDARD; PRT; 398 AA.
 ID LIPG_HUMAN
 AC P07098;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Triacylglycerol lipase, gastric precursor (EC 3.1.1.3) (Gastric
 DE lipase) (GL).
 GN LIPF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.


```

RX MEDLINE=87299724; PubMed=3304425;
RA Bodmer M.W., Angal S., Yarranton G.T., Harris T.J.R., Lyons A.,
RT King D.J., Pieroni G., Riviere C., Verger R., Lowe P.A.;
RT "Molecular cloning of a human gastric lipase and expression of the
RT enzyme in yeast.";
RL Biochim. Biophys. Acta 909:237-244(1987).
RN [2]
RP SEQUENCE OF 20-45.
RX MEDLINE=89325292; PubMed=27330332;
RA Bernbaeck S., Blackeberg L.;
RT "Human gastric lipase. The N-terminal tetrapeptide is essential for
RT lipid binding and lipase activity.";
RL Eur. J. Biochem. 182:495-499(1989).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX MEDLINE=99287897; PubMed=10358049;
RA Roussel A., Canaan S., Egloff M.P., Riviere M., Dupuis L., Verger R.,
RA Cambillau C.;
RT "Crystal structure of human gastric lipase and model of lysosomal
RT acid lipase, two lipolytic enzymes of medical interest.";
RL J. Biol. Chem. 274:16995-17002(1999).
CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O -> diacylglycerol + a
CC fatty acid anion.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE ABHDROLASE SUPERFAMILY. LIPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X05997; CAA29413.1;
CC EMBL: X05997; CAA29414.1; ALT_INIT.
CC EMBL: A01046; CAA00125.1;
CC EMBL: A12714; CAA01053.1;
CC PIR: S04942; S04942.
CC PIR: S07145; S07145.
CC PDB: 1HLG; 15-MAR-00.
CC Genew: HGNC:6622; LIPF.
CC MTM: 601980;
CC InterPro: IPR000073; Abhydrolase.
CC InterPro: IPR000734; Lipase.
CC Pfam: PF00561; abhydrolase; 1.
CC PROSITE: PS00120; LIPASE_SER; 1.
CC 3d-structure. 1
KW Hydrolase; Lipid degradation; Glycoprotein; Signal; Polymorphism;
KW 3d-structure. 1
FT SIGNAL 1 19
FT CHAIN 20 398 TRIACYLGLYCEROL LIPASE, GASTRIC.
FT ACT_SITE 172 172 CHARGE RELAY SYSTEM.
FT ACT_SITE 343 343 CHARGE RELAY SYSTEM.
FT ACT_SITE 372 372 CHARGE RELAY SYSTEM.
FT DISULFID 246 255
FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .).
FT VARIANT 161 161 T -> A (IN DBSNP: 814628).
FT FTID=VAR_011947.
FT SEQUENCE 398 AA; 45237 MW; CD3EE1621C014F0F CRC64;
Query Match 49.8%; Score 1119; DB 1; Length 398;
Best Local Similarity 52.2%; Pred. No. 1.6e-99;
Matches 210; Conservative 73; Mismatches 107; Indels 12; Gaps 2;
QY 17 MWLLILVAYFQKRVNNSVHPTKRAV-----DPEAFMNISEIIHQGYPCEEYEVAED 69
DB 1 MWLLLTWASLI-----SVLGTTHGLFKLHPGSPVMTNISQMTYNGYPNEEYEVVTD 55
QY 70 GYLISVNRIPRGLVOPKKTGRPVLLQHLGVGGASWNISLNPNSLGLFADAGFDVWM 129

```

```

Db 56 GYILEVNRIPYKKNKSGNTGQRPVFLQHGLLASATNWNISLNPNSLAFILADAGYDVL 115
QY 130 GNSRGNAWSKHKTLSDQDEFWAFSYDEMARDLPVINFILQKTQGEKIYVGVYSGQT 189
Db 116 GNSRGNTWARRNIYSPDSVEFWAFSEDEMAKYDLPATIDFIVKKTQKOLHVVHSGQT 175
QY 190 TMGFIASFSTMPELAQIKMYFALAPATVYKHAKSPTGKFLLLPDMIMKIGLFGKKEFLYQT 249
Db 176 TIGFIAFSTNPSLAKRIKTFYALAPVATVYKSLINKLRFVPSQLFKFIFGDKIFVPHN 235
QY 250 RFLROLVYILCGQVILDQICSNMLLLGGNTNNMNSRASYAAHTLACTSVQNTILHWS 309
Db 236 FFDQFLATEYCSREMLNLLCSNALFIICGDSKNFNTSLRDVYLSHNPACTSVQNMFWHT 295
QY 310 QAVNSGELRAFDSGETKNLEKNQPTPVRYRDMTPTAMMTGGODWLSNEDVKMLL 369
Db 296 QAVKSGKFAQYDNGSPVQNRHVDQSQPPYNTAMVPIAVNNGSKDLLADPDQVGLLL 355
QY 370 SEVTNLIYHKNPEMAHVDIFWGLDAPHRYNEIILHMQOEE 411
Db 356 PKLPNLIYHKEIPFYHNLDTIAMDAPQEVYNDIVSMISEDK 397
RESULT 8
LIP3_DROME
ID LIP3_DROME STANDARD; PRT; 394 AA.
AC Q46108;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lipase 3 precursor (EC 3.1.1.-).
GN LIP3 OR CG8823.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S.
RX MEDLINE=98227315; PubMed=9566193;
RA Pistillo D., Manzi A., Tino A., Pilo Boyl P., Graziani F., Malva C.;
RT "The Drosophila melanogaster lipase homologs: a gene family with
RT tissue and developmental specific expression.";
RL J. Mol. Biol. 276:877-885(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananathides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Douf L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

```

```
RA Merkulov G., Milshina N.V., Mobarrey C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- TISSUE SPECIFICITY: FAT BODY.
CC -1- DEVELOPMENTAL STAGE: ONLY AT LARVAL STAGES.
CC -1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.). ALSO SIMILAR TO
CC LEPIDOPTERAN EGG-SPECIFIC AND YOLK PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y14367; CA474737.1; ...
DR EMBL: AE003699; AAF54935.1; ...
DR FlyBase: FBgn0023495; Lip3.
DR InterPro: IPR000073; Abhydrolase.
DR InterPro: IPR000734; Lipase.
DR InterPro: IPR000379; Ser-estrase_site.
DR Pfam: PF00561; abhydrolase_1.
DR PROSITE: PS00120; LIPASE_SER; 1.
DR HydroLase: Lipid degradation; Signal: Glycoprotein.
FT SIGNAL 1 394
FT CHAIN 1 394
FT ACT_SITE 164 164
FT ACT_SITE 369 369
FT CARBOHYD 131 131
SQ SEQUENCE 394 AA; 44901 MW; A718D1D743673802 CRC64;
Query Match 29.5%; Score 663; DB 1; Length 394;
Best Local Similarity 37.08; Pred. No. 5.8e-50;
Matches 137; Conservative 80; Mismatches 139; Indels 14; Gaps 7;
QY 51 EIIHQGYPCPEYEVATEDGYILSYNRIPLGVQPK--KTGSRPVYLLQHLGLVGASNWI 108
DB 30 ERIEDGYPMERHEVYTSNDVILTMHRIP---YSPKTGESSNRPVAFELMHGWLSSDWW 86
QY 109 SNLPNNSLGLFIADAGFDVWNGNSRGNWSRKHKTLSDQDEFWAFSDDEWAFDLPVAVI 168
DB 87 LMGPSRLAYMLADAGYDVWNGNARGNTYSKAHKYTWPTYQIWFNFSNETGMDYDVPAMI 146
QY 169 NFILQTKGQKTYVYGYSGQTMGFATFSTPELAQKIMVAFALAPIATVYKHAQSPGTK- 227
DB 147 DIVLAKTGOQVQVYGHSGQITVYLVYVSERPEYNDKTKSAHLGLPAAYMGKMSPLTRA 206
QY 228 ----FLLPDMTKLFGKKEFLYQYRFLRLQLVYLCGVY-LDQICSNIMLLGFGFNWN 283
DB 207 FAPILQGPNAIVE-VCGSMFSPNSKFKODLGIEMCQATSPYADMCAANEIFLIGGYDTEQ 265
QY 284 MNMSRASVYAATLACTSVONILHWSQAVNSGELAPFWGSETKLEKCNQTPPYRVVR 343
DB 266 LDYELLEHIKATSPAGASVQNLHFCQYNSGKFRKFDY-TALRNPYEGSYFFPDYKIK 324
QY 344 DMTVPTAMWTGGDNLSPEDVKMLSEVTNLIYHKNTIP--EWAVHDFITGLDAPRMYN 401
DB 325 NAKAPVLIYGGANDWCDVSDVRKRLDELPLNMDLYLVFFERKWAHLDFITMGTEARVYD 384
```

```
QY 402 EIIHMQQEE 411
DB 385 EVLXMQMSYE 394
RESULT 9
LIP1_DROME STANDARD; PRT; 439 AA.
AC Q46107; J9VRR6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lipase 1 precursor (EC 3.1.1.-).
GN Lip1 OR G7279.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE OF 7-433 FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=98227315; PubMed=9566193;
RA Pastillo D., Manzi A., Tino A., Pilo Boyl P., Graziani F., Malva C.;
RT "The Drosophila melanogaster lipase homologs: a gene family with
RT tissue and developmental specific expression.";
RL J. Mol. Biol. 276:877-885(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Paulis B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Fodor C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck C.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel E.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mottel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarrey C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: COULD BE A DIGESTIVE ENZYME.
CC -1- TISSUE SPECIFICITY: OVARIES.
CC -1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.). ALSO SIMILAR TO
```

LEPIDOTERAN EGG-SPECIFIC AND YOLK PROTEINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: Y14366; CAA74736.1; -
EMBL: AF003629; RA752994.1; -
FlyBase: FBgn0023496; Lip1.
InterPro: IPR000073; Abhydrolase.
InterPro: IPR000734; Lipase.
InterPro: IPR000379; Ser_estrns_site.
Pfam: PF00561; abhydrolase_1.
PROSITE: PS00120; LIPASE_SER; FALSE_NEG.
Hydrolase; Lipid degradation; Signal; Glycoprotein.
FT SIGNAL 1 24
FT CHAIN 25 439
FT DOMAIN 30 44
FT CDS 197 197
FT ACT_SITE 197 197
FT ACT_SITE 393 393
FT CARBOHYD 124 124
FT CARBOHYD 151 151
FT CARBOHYD 346 346
FT CARBOHYD 379 379
FT CARBOHYD 426 426
FT CONFLICT 10 10
FT CONFLICT 213 213
FT CONFLICT 412 412
FT SEQUENCE 439 AA; 50660 MW; 9E3E20BEAE93E3F CRC64;
Quality Match 24.3%; Score 546; DB 1; Length 439;
Best Local Similarity 31.8%; Pred. No. 9.6e-40;
Matches 123; Conservative 83; Mismatches 157; Indels 24; Gaps 7;

QY 39 KAYDPEAFNMNISEIIHQGYPCBEYEYATEDGVYLSYNRIPRGLVQPKTGRSPVLLQH 98
DB 58 KNKQDSTLSVDKLIAGYGESEVHHVTTEDGYLTWHRI-----RKQGA-PPFLLQH 109
QY 99 GLVGGASWTSNLPNNSLGFILDADGVDVGMGSRGNMWSRKHKHTLSIDQDEFWAFSYDE 158
DB 110 GLVDSSAGFVYMGPNVSLAYLLADHNYDVWLGNARGNRYSRNHTTLPDDESKFWDSEWHE 169
QY 159 MARFDLPVINFILQKTGQEKIYVYVSGQFTMGFIASFTEPMLAQKIMYFALAPATV 218
DB 170 IGWYDLPAMIDHWLVTGTFPKLHYAGHSQGSCTFFVCMWRPAYNDKVRNMQALAPAVYA 229
QY 219 KHAKS-PGTRFLLPDMNIRKGLFGKREFLYQTRFLRQLVLYCGQVILDDQCSNIMLLG 277
DB 230 KETEDHPYIRALSIFYNSLVSGSIRENFGEFELCRMT-----ETERLCIEAVGIV 283
QY 278 GFNTNNMNSRASVYAHTLAGTSVQNILHWSQAVNSGELRAFDPMWGSETKNLEKCNQPT 337
DB 284 GRWNNEFNRMKFPVILGHYPAGVAAKQVKHFQIIKSGRFAPYSY-SSNKMQLYRDHLP 342
QY 338 VRYVRDVTYPTAMWTGGQDWLSNPEDYVKMLLSEVTLNLIYHKNP--EWAHVDFIWLGLDA 395
DB 343 PRYNSLSVTVPTFYVYSTNDLLCHPKDVKDSCMDLGNVTGKLYLPQKEFNHMDFLWAIDV 402
QY 396 PHRMYNIEIHLM-----QOEETNLSQ 416
DB 403 RKMLYRRMLQVLGVKPGCSPEEANSR 429

RESULT 10
ID TGL1 YEAST
AC P34163; STANDARD; PRT; 548 AA.
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)

```
RESULT 11
YD94_METJA          STANDARD;          PRT; 987 AA.
ID Q58789;
AC Q58789;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein MJ1394.
GN MJ1394.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcales; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT Jannaschii."
RL Science 273:1058-1073(1996).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO M.JANNASCHII MJ1393 AND A.FULGIDUS AF2028.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67579; AAB99404.1;
CC TIGR: MJ1394;
CC KW Hypothetical protein; Transmembrane; Complete proteome.
CC FT TRANSMEM 12 32
CC FT TRANSMEM 958 978
CC FT TRANSMEM 987 AA; 112360 MW; D1E628FF28CA86D CRC64;
CC SQ SEQUENCE 987 AA; 112360 MW; D1E628FF28CA86D CRC64;
CC -----
Query Match
Best Local Similarity 23.1%; Pred. No. 0.0079;
Matches 105; Conservative 56; Mismatches 149; Indels 144; Gaps 29;
QY 3 ETLSRQIVSHRMEMWLLILVAYMFQR-----NVNSVHMPTKAVDPEAFMN-----48
DB 340 ETVDTVPKANKMSEVNLVLYLNKLVPGENVNVNVIKYLNSGQFYDNFENWGTWQ 399
QY 49 -----ISEIIHQG-YCEEYEVEATEDGYILSVNIRPG--LVOPKKTGSRPVVLLQH 98
DB 400 YKNGIVQWMSIQSHSNYSLEKYGIST-----SLANDPNGGYKLLPKEIG-RDVI--S 450
QY 99 GLVGGASNWIENLPPNNSLGFILADAGFDVWGMNSRGNASKHKT--LSIDQ-----148
DB 451 GWYRPSNW-GGFPIDRIG--LEDENFD-----GYSEVNHYSNYSIDRTNNGNPT 500
QY 149 ---DEFWAFSDYEMARFDLPVINFILQKTQGEKYYVYGSGCTMGTFATFSTMPELAQ 205
DB 501 ISPEVYNPPDEWYFEL-----KI-----YSNCT---ITFST-----531
QY 206 IKWYFALAPI-ATVKKAKSPGTF-----LLLPDMWIKG-----LFGKKEFLYQTR 250
DB 532 ---YYQNGSLAATVSTIDNTYTKFDVRVHGGVYVYVDDLEVNSKNFDYGDKNWYK--586
QY 251 FLRLVYILGGVILQICSNIMLLGG-----FNTNMNMSRASVYAAHTL-----AG 299
DB 587 -----EITSANSEGTAVLFDGDFYKDKDINTSLN---AINWNTITLWNSDSA 632
```

```
QY 300 TSVQNIL-HWSQAVNSGELRAFDWGSSEKTKNLEKCNQPTPVR--YVRVDMTVPVTAMWTGGQ 356
DB 633 TLVENVLGNYSYSERDNLAKYGFILKFNNGTNTNTSIKGVYASGSYSISTDHGTTGE 692
QY 357 ---DWLSN---PEDVKMLLSEVNTLIYHKNIPEWA 385
DB 693 INIWIENVTFRKNDAKSYSNFLTNL-----NI--WA 720
```

```
RESULT 12
ORC2_CAEEL
ID ORC2_CAEEL          STANDARD;          PRT; 430 AA.
AC Q21037;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Origin recognition complex subunit 2 (Ceorc2).
GN ORC-2 OR F59E10.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=96099401; PubMed=7502077;
RA Gavin K.A., Hidaka M., Stillman B.;
RT "Conserved initiator proteins in eukaryotes."
RL Science 270:1667-1671(1995).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Swinburne J.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
REVIEWS.
RA Jones S.J.M.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF THE ORIGIN RECOGNITION COMPLEX (ORC) THAT
CC BINDS ORIGINS OF REPLICATION. IT HAS A ROLE IN BOTH CHROMOSOMAL
CC REPLICATION AND MATING TYPE TRANSCRIPTIONAL SILENCING. BINDS TO
CC THE ARS CONSENSUS SEQUENCE (ACS) OF ORIGINS OF REPLICATION IN AN
CC ATP-DEPENDENT MANNER.
CC -!- SUBUNIT: ORC IS COMPOSED OF SIX SUBUNITS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE ORC2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U40270; AAC46954.1;
CC EMBL: Z36949; CAA85415.1;
CC WormPep: F59E10.1; CB11492.
CC DNA replication; Nuclear protein.
CC SQ SEQUENCE 430 AA; 49319 MW; 54FC086BB4AD9670 CRC64;
```

```
Query Match
Best Local Similarity 4.7%; Score 105; DB 1; Length 430;
Matches 61; Conservative 42; Mismatches 87; Indels 88; Gaps 13;
QY 85 PKTGRPVVLLQHLGVGGASNWIENLPPN-----SLGF-----ILADAGFDW--MGN 131
DB 25 PEKESR-----QKTKNGKENASRLQSLNEEDLEQLGFEDTVTSQAIEYFQWKG 78
QY 132 S-----RGNASWRKHK-----TSLIDQDEFWAFSDYEMARFDLPVINFILQKTQGEK 181
DB 79 SASERNNAKSRGRAGNAGNTEIEDEDEDEISNAITDFTKCDLPGLRNYTKKDNTEFEK 138
```

QY 182 YVGSQGTMGFIATFSTPELAQKIKMYFALAPATVTKHAKSPGTFKLLPDMNKGFLG 241
 DB 139 RLEHLADNDFG-----KWKLYLAAG-----FNILLHGVGS 168
 QY 242 KKEFLYQRFRLQRLVYLCGOVILDOICSNIMLLGGFNTN---NNMNSRASVYAHTLA 298
 DB 169 KRDVL--TEFENELSDYYMYRVDARKDLNKKVLLGALNENKKNVGRG-----217
 QY 299 GTSVQNLHWSQAV-----NSGEL-----RAFDWGS 325
 DB 218 ---QSTISWARSIRKKNNSOOLILIIDNIEAPDWRSD 251

RESULT 13
 ID PAFA_CAVPO STANDARD: PRT: 436 AA.
 AC P70683;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Platelet-activating factor acetylhydrolase precursor (EC 3.1.1.47)
 DE (PAF acetylhydrolase) (PAF 2-acetylhydrolase) (LDL-associated
 DE phospholipase A2) (LDL-PAF(2)) (2-acetyl-1-alkylglycerophosphocholine
 DE esterase) (1-alkyl-2-acetyl-1-alkylglycerophosphocholine
 GN PLAZG7 OR PAFAH.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN-Hartley; TISSUE=Liver;
 RX MEDLINE=97103479; PubMed=8947850;
 RA Karasawa K., Kuge O., Kawasaki K., Nishijima M., Nakano Y., Tomita M.,
 RA Yokoyama K., Setaka M., Nojima S.;
 RT Cloning, expression and characterization of plasma
 RT platelet-activating factor-acetylhydrolase from guinea pig.";
 RL J. Biochem. 120:838-844(1996).
 CC -1- FUNCTION: MODULATES THE ACTION OF PLATELET-ACTIVATING FACTOR (PAF)
 CC BY HYDROLYZING THE SN-2 ESTER BOND TO YIELD THE BIOLOGICALLY
 CC INACTIVE LYSO-PAF. HAS A SPECIFICITY FOR SUBSTRATES WITH A SHORT
 CC RESIDUE AT THE SN-2 POSITION. IT IS INACTIVE AGAINST LONG-CHAIN
 CC PHOSPHOLIPIDS.
 CC -1- CATALYTIC ACTIVITY: 2-acetyl-1-alkyl-sn-glycero-3-phosphocholine +
 CC H(2)O = 1-alkyl-sn-glycero-3-phosphocholine + acetate.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL: D67037; BAAL1054.1;
 DR InterPro: IPR000734; Lipase.
 DR InterPro: IPR005065; PAF-AH_P-II.
 DR Pfam: PF03403; PAF-AH_P-II; 1.
 DR PROSITE: PS00120; LIPASE_SER; 1.
 KW Hydrolase; Lipid degradation; Glycoprotein; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 436
 FT PLATELET-ACTIVATING FACTOR
 FT ACETYLHYDROLASE.
 FT ACT_SITE 271
 FT CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 294
 FT CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 349
 FT CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CARBOHYD 76
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 200
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 324
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 436 AA; 49062 MW; C359D96E392FFE11 CRC64;
 Query Match 4.5%; Score 101; DB 1; Length 436;
 Best Local Similarity 20.4%; Pred. No. 0.38;
 Matches 85; Conservative 60; Mismatches 149; Indels 122; Gaps 21;

QY 21 ILVAFMFORNVNSVMPKAVD--PEAFMNISEIIOHQGPCEEY-EVATED-----GY 71
 DB 70 LMSGYTNQSSFLRLYPSQNDNDFDALW-----IPNEEYFQGLTETLGLASSFLGK 119
 QY 72 ILSV---NRIPRGLVQPKKTSR--PVYLLOHGLVGGASNMISNLPNNSLGFILADAGFD 126
 DB 120 LKLLYGSVKVPKAKWNSPLKTKGKYLIFSHGL--GAFRSI-----YSAIGIELASHGFI 173
 QY 127 V-----WMGNSRGNASRKRHKHTLSIDQDE-----FW 152
 DB 174 VAAVEHRDESAAATYFYFDAPAAESGRNSRWYKVGNLTEERKRQLRQGECSQALSW 233
 QY 153 AFSYDEMARPDLPAVINFILOK-----TQGEKIYVYVYSGQTMGFIATFSTPELAQKIKM 208
 DB 234 LLSIDGEGEPKVNLDNFDIOQLKGLSDRSKVAIIHGSFG--GATVIQTLSE--DQRFRC 289
 QY 209 YFAL---APIATVKHAKSPGTFKLLPDMNKGFLGKK-BELYOTRFRLQVL-----256
 DB 290 GIALDPWMPVGEDVHSKIPOPLFFINSEYFQSANDTKIEKFYQPKERKMAIVKGSVH 349
 QY 257 -----IVLCGQVILDOICSNIMLLGGFNTNMNN---SRASVYAHTLACTSVQNLH 307
 DB 350 HNFVDFTFAGKII-----GQMSLKGKIDSEVADMLINKASLAFLOKYLGLD-KNFQD 402
 QY 308 WSOAVNSGELRAFDWGSSEKLEKCNQPTPVRYRVDMTVPTAMTGGQDWLSNPE 363
 DB 403 WNSLME-----GDDENL-----IPEETIPTTMOSSGTGEORND 436

RESULT 14
 ID Y193_HAEIN STANDARD: PRT: 287 AA.
 AC Q57427; O05013;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative esterase/lipase H10193 (EC 3.1.1.-).
 GN H10193.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]

SEQUENCE FROM N.A.
 RP STRAIN-Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kervatage A.R., Sult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RL Science 269:495-512(1995).
 CC -1- SIMILARITY: BELONGS TO A FAMILY OF ESTERASES THAT GROUPS TOGETHER
 CC PSEUDOMONA TROPINSTERASE, DMPD, TODF AND XYL.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

```

Query Match      4.3%; Score 95.5; DB 1; Length 1259;
Best Local Similarity 19.1%; Pred. NO. 4.8;
Matches 50; Conservative .67; Mismatches 170; Indels 129; Gaps 23;

QY 28 ORNVNSVHMTKAVDPEAFNII---SEIIHQGQYPCPEYE----- 64
      | | | | | | | | | | | | | | | | | | | | | |
Db 442 QEVESLNRDITSEIEAIINSLPNKKSPGEGFTAEFYQRYKEELVFLKLFQSIKE 501
      | | | | | | | | | | | | | | | | | | | | | |
QY 65 -VATEDGYILSVNRIPR-GLVQPKKTKTSRPVWL-----LQHLVGGSANWISNL-PNN 114
      | | | | | | | | | | | | | | | | | | | | | |
Db 502 GIIPNSVEYASIIILPKPGDRTTKKKNFRPISLNMNTDAKTLNKILANOIOOHTKKIITHD 561
      | | | | | | | | | | | | | | | | | | | | | |

```

```

115 SLAFILADAGFDVMMGNSR-----GNAWSRKHKTLSDIQDFWFAVSDEM----- 159
      :||:|
Db    562 QVGFIPAMQG---WFNIRKSGINIHOINRKTDMHMIISIDAEK-----AFDKIQOOPMLK 614
      :||:|
Qy    160 -----ARFDPAVINFILOKTGOEKIYVYVYSGQTTMGFIASFMPELA 203
      :||:|
Db    615 PLHLKLGIDGYLKIIRAIYDKPTA-NIIL-NGO-KLEAPPLKTGTGRQGCPLSPLENNIV 670
      :||:|
Qy    204 QKZKMYEPALAPIATVHKHAKSPGKFLLLPDMWTIKGFKKEFLYQTRFLRQLVYILCGQV 263
      :||:|
Db    671 LEV-----LA--RAIROEKE-----IKGIOLGKEEVKLSLFADDIMVYLENPI 711
      :||:|
Qy    264 ILDQICSNIMLLGGFN-----TNNMMNSRASVYAAHTLAGTSVONILHWSQAVNSGELR 318
      :||:|
Db    712 VSAQ---NLLKLISFSGVSKYKINVQKSAFLYTNNRQTESQIMSELPFTTISKRIKYL 768
      :||:|
Qy    319 AFDWGSETKNLEKXCNOPTPYRVYRVDWT-----VPTAMWTGGQDWLSNPEDVKMLLSEVT 373
      :||:|
Db    769 GIQLTRQVOKDLFENY-KPLLNETIKEDTNKWNIPCS-----WVGRIINIVKMAI--LP 818
      :||:|
Qy    374 NLIVHKN-IP-----ENAHVDVFINGLDAPH 397
      :||:|
Db    819 KVIYFRNAIPKLPMTFFTELEKTTLAFIWNOKRAH 854
      :||:|

```

Search completed: February. 3, 2003, 13:49:16
Job time : 17 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: February 3, 2003, 13:47:24 ; Search time 35 Seconds
(without alignments)
2490.227 Million cell updates/sec

Title: US-09-333-159-47
Perfect score: 2247
Sequence: 1 MLSTLSRQWIVSRHEWMLL.....IHLMQOETNLSQRCRAVL 423

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1991	88.6	374	4	Q96LG2
2	1134.5	50.5	395	11	Q9D798
3	1134.5	50.5	395	11	Q9D6X0
4	1134.5	50.5	395	11	Q9D6Q2
5	1134.5	50.5	395	11	Q9D6G9
6	1133.5	50.4	395	11	Q9D6P3
7	1133.5	50.4	395	11	Q9D6P8
8	1131.5	50.4	395	11	Q9D6P7
9	1129.5	50.4	395	11	Q9D6L9
10	1128.5	50.3	395	11	Q9D6T5
11	1127.5	50.2	395	11	Q9D767
12	1126.5	50.1	395	11	Q9D7C5
13	1126.5	50.1	395	11	Q9D766
14	1125.5	50.1	395	11	Q9D6L1
15	1124.5	50.0	395	11	Q9D6Q3
16	1124.5	50.0	395	11	Q9D760
					Q9D6S5

17	1122.5	50.0	395	11	Q9D796	Q9D796 mus musculu
18	1119.5	49.8	395	11	Q9D6N8	Q9D6N8 mus musculu
19	1116.5	49.7	395	11	Q9D6Q2	Q9D6Q2 mus musculu
20	1115.5	49.6	395	11	Q9D6G9	Q9D6G9 mus musculu
21	755.5	33.6	684	5	Q95XV1	Q95XV1 caenorhabdi
22	752.5	33.5	403	5	Q61866	Q61866 caenorhabdi
23	751	33.4	411	5	Q20449	Q20449 caenorhabdi
24	730	32.5	405	5	Q93789	Q93789 caenorhabdi
25	716	31.9	404	5	Q16956	Q16956 caenorhabdi
26	715	31.8	169	11	Q9D2L7	Q9D2L7 mus musculu
27	692	30.8	411	5	Q94252	Q94252 caenorhabdi
28	668.5	29.8	351	5	Q95X33	Q95X33 caenorhabdi
29	655.5	29.2	426	5	Q17766	Q17766 caenorhabdi
30	655.5	29.2	434	5	Q9VKT9	Q9VKT9 drosophila
31	639	28.4	456	5	Q9VKS5	Q9VKS5 drosophila
32	636	28.3	398	5	Q9V796	Q9V796 drosophila
33	625.5	27.8	616	5	Q77107	Q77107 plodia inte
34	588.5	26.2	838	5	Q9V8K6	Q9V8K6 drosophila
35	574.5	25.6	559	5	Q17219	Q17219 bombyx mori
36	567.5	25.3	457	5	Q9VKT2	Q9VKT2 drosophila
37	567.5	25.3	457	5	Q8T3X7	Q8T3X7 drosophila
38	559	24.9	399	5	Q9VPE9	Q9VPE9 drosophila
39	535	23.8	416	5	Q9VQ05	Q9VQ05 drosophila
40	529.5	23.6	406	5	Q9VKT7	Q9VKT7 drosophila
41	527	23.5	504	5	Q94568	Q94568 galliera me
42	505	22.5	443	5	P78898	P78898 schizosacch
43	502	22.3	435	5	Q9VG46	Q9VG46 drosophila
44	489	21.8	457	5	Q95U37	Q95U37 drosophila
45	488	21.7	355	5	Q9VKT0	Q9VKT0 drosophila

ALIGNMENTS

RESULT 1

Q96LG2 ID Q96LG2 PRELIMINARY; PRT; 374 AA.
AC Q96LG2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE BA30415.1 (Novel lipase) (Fragment).
GN BA30415.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bray-Allen S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL353113; CAC78754.1; -
DR InterPro: IPR000073; Abhydrolase.
DR InterPro: IPR000734; Lipase.
DR InterPro: IPR000379; Ser_estrs_site.
DR Pfam: PF00561; abhydrolase; 1.
DR PROSITE, PS00120; LIPASE_SER; UNKNOWN_1.
FT NON_TER
SQ SEQUENCE 374 AA; 42371 MW; 5E7220A889437337 CRC64;

Query Match 88.6%; Score 1991; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. No. 9e-167;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	50	SEIIIOHQYPCPEEYEVATEDGYILSVNRPGLVQPKKTSRPVYLQHLGLVGGASNWIS	109
Db	1	SEIIIOHQYPCPEEYEVATEDGYILSVNRPGLVQPKKTSRPVYLQHLGLVGGASNWIS	60
Qy	110	NLPNNSLGFIADAGFDVWGMGNSRGNVNSRKHKTLSIDODEFWAFSYDEMAFDLPVAVIN	169
Db	61	NLPNNSLGFIADAGFDVWGMGNSRGNVNSRKHKTLSIDODEFWAFSYDEMAFDLPVAVIN	120
Qy	170	FIEQKTGQEKIYYVSGQTTMGTFIAFTMPLEAQIKRMYPALATIATYKHAKSPTGKFL	239

```
Db 121 FILQKGTGKYYVYSGTGMGFIATFSTPELAQKIKMYFALAPIATVHKASPGTRFL 180
Qy 230 LLPDMWIKLFGKKEFLYQTRFLQOLYVILCGOVLDOICSNIMLLGGENTNNMNSRA 289
Db 181 LLPDMWIKLFGKKEFLYQTRFLQOLYVILCGOVLDOICSNIMLLGGENTNNMNSRA 240
Qy 290 SVYAHTLAGTSVONILHWSQAVNSGELRAFDWGSSEKTKLEKCNQPTPVRYVRDVTPT 349
Db 241 SVYAHTLAGTSVONILHWSQAVNSGELRAFDWGSSEKTKLEKCNQPTPVRYVRDVTPT 300
Qy 350 AMWTGGQDWSNPEDVKMLLSEVNTLYHKNIPEVAHVDFIWLGLDAPHMYNEIHLMOQ 409
Db 301 AMWTGGQDWSNPEDVKMLLSEVNTLYHKNIPEVAHVDFIWLGLDAPHMYNEIHLMOQ 360
Qy 410 EETNLSQGRCAVL 423
Db 361 EETNLSQGRCAVL 374

RESULT 2
Q9D798 ID Q9D798 PRELIMINARY; PRT; 395 AA.
AC Q9D798;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE 2310051B21RIK protein.
GN 2310051B21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21083660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK009431; BAB26283.1;
DR MGD; MGI:1914967; 2310051B21RIK.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44603 MW; D3FDB9B6FA671E3E CRC64;

Query Match 50.5%; Score 1134.5; DB 11; Length 395;
Best Local Similarity 53.5%; Pred. No. 2.2e-91;
Matches 213; Conservative 70; Mismatches 108; Indels 7; Gaps 3;

Qy 17 MWLLTLVAYMFQ-RNVNSVHPTKAVDPAFNPINSEITQHOGYPCPEEYEVATEDGYILSV 75
Db 1 MWLLLVTVLFAFGAGHGLFKGLPKNPANMNVSQMITYNGYPSPEYEVATEDGYILGV 60
```

Query Match 50.5%; Score 1134.5; DB 11; Length 395;

Best Local Similarity 53.5%; Pred. No. 2.2e-91;
Matches 213; Conservative 70; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLILVAYFQ-RNVASVHMPTRKAVDPEAFMNISEIIHQGYPCPEEYEVATEDGYILSV 75
DB 1 MWLLVTSVLSAFGAGHGLFGKLPKPNPEANNVSMQITWGYPSEEVTEDEGYILGV 60

QY 76 NRIPRGLVQPKTSRPPVLLQHLGVGASNWSNLPNSLGLFILDADGFDVWNGSRGN 135
DB 61 YRIPYKGNSENICKRPVAYLQHLGVASATNWTNLPNSLAFILADAGYDVLGNSRGN 120

QY 136 AWSKHKTLSTLDQDEFWAFSDENARFDLPVINFILQKTOEKIYVYVSGTGMGFA 195
DB 121 TWSRKNVYSPDSVEFWAFSDENARFDLPVINFILQKTOEKIYVYVSGTGMGFA 180

QY 196 FSTMPLEAKTKMYFALAPATVATVHKASPGTKFLLPDMIMKGLFGKKEFL---YOTRFL 252
DB 181 FSTNPALAKKIKRYALAPVATVYTESPFKKISLIPKFLKLVIFGNKMFPHNYLDQFL 240

QY 253 RQLVYLCGGVILQDQICSNMILLGGFNTNNMNSRASVYAAHTLAGTSVQNLHWSQAV 312
DB 241 GTEV---CSRELLDLCSNALFIFCGFDKKNLVSREFDVLGHNPACTSTQDLFHWQA 297

QY 313 NSGELRAFDMGSETKNLEKCNQPTPVRYVRDMVPTAMTGGQDMLNPEDEVKMLSEV 372
DB 298 KSGKLOAYNMGSPQLQNLHYNQKTPPYDVVSAMTVPIAVNNGGHDILADPODVANLLPKL 357

QY 373 TNLVYHKNIPEMAHVDFTWGLDAPHRMYNEIHLMOQE 410
DB 358 PNLVYHKELPYNLHLDIFWAMDAPQEVYNEIVTMMAED 395

RESULT 4
Q9D606 PRELIMINARY; PRT; 395 AA.

AC Q9D606;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:2310067K20, full insert sequence.
GN 2310051B21R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyszynski A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010093; BAB26697.1;
DR MGD; MGI:1914967; 2310051B21R1K.
DR InterPro; IPR000073; Abhydrolase.

DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estr_ site.
DR Pfam; PF00561; abhydrolase_1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44623 MW; 245055F5E7FF07C9 CRC64;

Query Match 50.5%; Score 1134.5; DB 11; Length 395;
Best Local Similarity 53.8%; Pred. No. 2.2e-91;
Matches 214; Conservative 69; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLILVAYFQ-RNVASVHMPTRKAVDPEAFMNISEIIHQGYPCPEEYEVATEDGYILSV 75
DB 1 MWLLVTSVLSAFGAGHGLFGKLPKPNPEANNVSMQITWGYPSEEVTEDEGYILGV 60

QY 76 NRIPRGLVQPKTSRPPVLLQHLGVGASNWSNLPNSLGLFILDADGFDVWNGSRGN 135
DB 61 YRIPYKGNSENICKRPVAYLQHLGVASATNWTNLPNSLAFILADAGYDVLGNSRGN 120

QY 136 AWSKHKTLSTLDQDEFWAFSDENARFDLPVINFILQKTOEKIYVYVSGTGMGFA 195
DB 121 TWSRKNVYSPDSVEFWAFSDENARFDLPVINFILQKTOEKIYVYVSGTGMGFA 180

QY 196 FSTMPLEAKTKMYFALAPATVATVHKASPGTKFLLPDMIMKGLFGKKEFL---YOTRFL 252
DB 181 FSTNPALAKKIKRYALAPVATVYTESPFKKISLIPKFLKLVIFGNKMFPHNYLDQFL 240

QY 253 RQLVYLCGGVILQDQICSNMILLGGFNTNNMNSRASVYAAHTLAGTSVQNLHWSQAV 312
DB 241 GTEV---CSRELLDLCSNALFIFCGFDKKNLVSREFDVLGHNPACTSTQDLFHWQA 297

QY 313 NSGELRAFDMGSETKNLEKCNQPTPVRYVRDMVPTAMTGGQDMLNPEDEVKMLSEV 372
DB 298 KSGKLOAYNMGSPQLQNLHYNQKTPPYDVVSAMTVPIAVNNGGHDILADPODVANLLPKL 357

QY 373 TNLVYHKNIPEMAHVDFTWGLDAPHRMYNEIHLMOQE 410
DB 358 PNLVYHKELPYNLHLDIFWAMDAPQEVYNEIVTMMAED 395

RESULT 5
Q9D6P3 PRELIMINARY; PRT; 395 AA.

AC Q9D6P3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:2310069P19, full insert sequence.
GN 2310051B21R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyszynski A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010093; BAB26697.1;
DR MGD; MGI:1914967; 2310051B21R1K.
DR InterPro; IPR000073; Abhydrolase.

```

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010139; BAB26725.1; -
DR MGI; MGI:1914967; 2310051B21Rik.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000734; Lipase.
DR Pfam; PF00561; abhydrolase_1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44651 MW; E48A73CCCBFD359D CRC64;

Query Match 50.5%; Score 1134.5; DB 11; Length 395;
Best Local Similarity 53.8%; Pred. No. 2.2e-91;
Matches 214; Conservative 69; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFO-RNVNSVHMPTKAVDPAFNISEIIOHQGYPCPEEYEVATEDGYLSV 75
DB 1 MWLLLVTSVLSAFGGAHGLFGKLGPKNPANMVNSOMITYWGPSEEEYEVATEDGYLGV 60

QY 76 NRIPRGLVQPKTGSRRPVYLLQHLGVGASNINLNNLSGLFILADAGFDVWNGNSRN 135
DB 61 YRIPYKKNSENIGRPVAYLQHLGIASATNITLNNLSLAFILADAGYDVLGNSRN 120

QY 136 AWSRKHKTLSDODEFWAFSDYEMARFDLPVINFILQKTQGEKIYYVGSQGTMGFTA 195
DB 121 TWSRKNVYSPDSVEWAFSFEDEMAKYDLPATIDFIVQKTQGEKIHYVGHSGTIGFTA 180

QY 196 FSTMPELAQIKMYFALAPIATVYKAKSPGTFKLLPDMWIKGLFGKKEFL---YQTRFL 252
DB 181 FSTNPAKAKIKRFYALAPIATVYKTESPFKISLIPKFLKLVIFGNKMFPHNYLQDFL 240

QY 253 ROLVYLQGVQILDOICSNIMLLGGFNINNNMNSRASVYAAHTLAGTSVQNLHWSQAV 312
DB 241 GTEV---CSRELLDLCSNALFICGDFKKNLVSRFDVYLGHNPACTQTDLFHWQA 297

QY 313 NSGELRAFDWGSSETNLEKCNQPTPVRYRVDMTVPYVMTQGDWLSNPEDVKMLLSEV 372
DB 298 KSGKLQAYNWGSPQLQNLHYNQKTPPYDVSANTVPIAVWNGHDLADPQDVAMLLPKL 357

QY 373 TNLVYHKNPEWAHVDFTWGLDAPHRMYNEIHLMOOE 410
DB 358 PNLLYHKEILPYNHLDFIWADAPQEVYNEIVTMAED 395

RESULT 6
Q9C9P8 PRELIMINARY; PRT; 395 AA.
AC Q9C9P8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE 2310051B21Rik protein.
GN 2310051B21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10990;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawaji J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Konno H., Adachi J., Fukuda S.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glisic C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

```

```

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyono-Oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010116; BAB26711.1; -
DR EMBL; AK009300; BAB26201.1; -
DR MGI; MGI:1914957; 2310051B21Rik.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000734; Lipase.
DR Pfam; PF00561; abhydrolase_1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44579 MW; D2996865EA671D34 CRC64;

Query Match 50.4%; Score 1133.5; DB 11; Length 395;
Best Local Similarity 53.5%; Pred. No. 2.7e-91;
Matches 213; Conservative 70; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFO-RNVNSVHMPTKAVDPAFNISEIIOHQGYPCPEEYEVATEDGYLSV 75
DB 1 MWLLLVTSVLSAFGGAHGLFGKLGPKNPANMVNSOMITYWGPSEEEYEVATEDGYLGV 60

QY 76 NRIPRGLVQPKTGSRRPVYLLQHLGVGASNINLNNLSGLFILADAGFDVWNGNSRN 135
DB 61 YRIPYKKNSENIGRPVAYLQHLGIASATNITLNNLSLAFILADAGYDVLGNSRN 120

QY 136 AWSRKHKTLSDODEFWAFSDYEMARFDLPVINFILQKTQGEKIYYVGSQGTMGFTA 195
DB 121 TWSRKNVYSPDSVEWAFSFEDEMAKYDLPATIDFIVQKTQGEKIHYVGHSGTIGFTA 180

QY 196 FSTMPELAQIKMYFALAPIATVYKAKSPGTFKLLPDMWIKGLFGKKEFL---YQTRFL 252
DB 181 FSTNPAKAKIKRFYALAPIATVYKTESPFKISLIPKFLKLVIFGNKMFPHNYLQDFL 240

QY 253 ROLVYLQGVQILDOICSNIMLLGGFNINNNMNSRASVYAAHTLAGTSVQNLHWSQAV 312
DB 241 GTEV---CSRELLDLCSNALFICGDFKKNLVSRFDVYLGHNPACTQTDLFHWQA 297

QY 313 NSGELRAFDWGSSETNLEKCNQPTPVRYRVDMTVPYVMTQGDWLSNPEDVKMLLSEV 372
DB 298 KSGKLQAYNWGSPQLQNLHYNQKTPPYDVSANTVPIAVWNGHDLADPQDVAMLLPKL 357

QY 373 TNLVYHKNPEWAHVDFTWGLDAPHRMYNEIHLMOOE 410
DB 358 PNLLYHKEILPYNHLDFIWADAPQEVYNEIVTMAED 395

RESULT 7
Q9C9P7 PRELIMINARY; PRT; 395 AA.
AC Q9C9P7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE 2310051B21Rik protein.
GN 2310051B21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10990;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawaji J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Konno H., Adachi J., Fukuda S.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glisic C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

```

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fieischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK010236; BAB26787.1; -;
DR EMBL: AK009413; BAB26272.1; -;
DR EMBL: AK009428; BAB26280.1; -;
DR EMBL: AK009459; BAB26300.1; -;
DR EMBL: AK009473; BAB26312.1; -;
DR EMBL: AK009474; BAB26313.1; -;
DR EMBL: AK009479; BAB26316.1; -;
DR EMBL: AK009523; BAB26338.1; -;
DR EMBL: AK009525; BAB26339.1; -;
DR EMBL: AK009546; BAB26352.1; -;
DR EMBL: AK009571; BAB26358.1; -;
DR EMBL: AK009573; BAB26370.1; -;
DR EMBL: AK009729; BAB26466.1; -;
DR EMBL: AK009773; BAB26495.1; -;
DR EMBL: AK010019; BAB26647.1; -;
DR EMBL: AK010035; BAB26656.1; -;
DR EMBL: AK010058; BAB26673.1; -;
DR EMBL: AK010061; BAB26675.1; -;
DR EMBL: AK010124; BAB26715.1; -;
DR EMBL: AK010125; BAB26716.1; -;
DR MGD: MGI:1914967; 2310051B21Rik.
DR InterPro: IPR000073; Abhydrolase.
DR InterPro: IPR000734; Lipase.
DR Pfam: PF00561; abhydrolase; 1.
DR PROSITE: PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44637 MW; D3F96B65EA671E34 CRC64;

Query Match 50.4%; Score 1133.5; DB 11; Length 395;
Best Local Similarity 53.5%; Pred. No. 2.7e-91;
Matches 213; Conservative 70; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFQ-RNVNSVHMPKAVDPEAFPMNISEIIHQGYPCPEYEYVATEDGYILSV 75
Db 1 MWLLLVTSVLSAFSGAGHGLFGKLGPKNPEANNVSMITYWGYPSEYEVYEDGYILGV 60

QY 76 NRIPRGLVQPKKTSRPPVLLQHLGVGASNWSINLPNSLGLFIADAGFDVWNGSRGN 135
Db 61 YRIPYKGNSENIGKRPVAVLQHLGLIASATNWTINLPNSLAFILADAGYDVLNLSRGN 120

QY 136 AWSKKHTLSIDQDEFWAFSDENARFDLPVINFILQKTQGEKIYVGYSGQTMGFIA 195
Db 121 TWSRKNVYSPDSVEFWAFSEDEMAKYDLPATIDFIVQKTQGEKIHYVGHSGQTIGFIA 180

QY 196 FSTMPLEAOKIMYFALAPATVATVKHAKSPGTFKLLPDMIMKGLFGKKEFL---YQTRFL 252
Db 181 FSTNPALAKKIKRYPALAPATVATVKYTESPPKISLIPKELLKYIFGNKMFPHNYLDQFL 240

QY 253 RQLVIYLCGGVILQDQICSNIMLLGGFTNNMNSRVSVAHTLAGTSVQNIHLHWSQAV 312
Db 241 GTEY---CSRLLDLLCSNALFIFCGFDKKNLVSRFDVYLGHNPAGTSTQDLFHWQA 297

QY 313 NSGELRAFDMGSETKNEKCNQTPPVRYVRDNTVPTAMTGGODWLSNEDYKMLLSEV 372
Db 298 KSGKLQAINWGSPLQNNLHYNQKTPPYIDVSAMTVPVAVWNGGHDILADFDQVAMLLPKL 357

QY 373 TNLIIYHKNIPEWAVDFIWLGDAPHRMYNEIHLMOOE 410
Db 358 PNLIIYHKKEILPYNHLDFIWMADAPQEVYNEIVTWMAD 395

RESULT 8
Q9D6L9
ID Q9D6L9 PRELIMINARY; PRT; 395 AA.
AC Q9D6L9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:23100761i3, full insert sequence.
GN 2310051B21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamahaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fieischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK010203; BAB26766.1; -;
DR MGD: MGI:1914967; 2310051B21Rik.
DR InterPro: IPR000073; Abhydrolase.
DR InterPro: IPR000734; Lipase.
DR InterPro: IPR000379; Ser_estrs_site.
DR Pfam: PF00561; abhydrolase; 1.
DR PROSITE: PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44671 MW; B8936162510AA55C CRC64;

Query Match 50.4%; Score 1131.5; DB 11; Length 395;
Best Local Similarity 53.5%; Pred. No. 4e-91;
Matches 213; Conservative 69; Mismatches 109; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFQ-RNVNSVHMPKAVDPEAFPMNISEIIHQGYPCPEYEYVATEDGYILSV 75
Db 1 MWLLLVTSVLSAFSGAGHGLFGKLGPKNPEANNVSMITYWGYPSEYEVYEDGYILGV 60

QY 76 NRIPRGLVQPKKTSRPPVLLQHLGVGASNWSINLPNSLGLFIADAGFDVWNGSRGN 135
Db 61 YRIPYKGNSENIGKRPVAVLQHLGLIASATNWTINLPNSLAFILADAGYDVLNLSRGN 120

QY 136 AWSKKHTLSIDQDEFWAFSDENARFDLPVINFILQKTQGEKIYVGYSGQTMGFIA 195
Db 121 TWSRKNVYSPDSVEFWAFSEDEMAKYDLPATIDFIVQKTQGEKIHYVGHSGQTIGFIA 180

QY 196 FSTMPLEAOKIMYFALAPATVATVKHAKSPGTFKLLPDMIMKGLFGKKEFL---YQTRFL 252
Db 181 FSTNPALAKKIKRYPALAPATVATVKYTESPPKISLIPKELLKYIFGNKMFPHNYLDQFL 240

QY 253 RQLVIYLCGGVILQDQICSNIMLLGGFTNNMNSRVSVAHTLAGTSVQNIHLHWSQAV 312

QY 76 NRIPRGLVOPKTKGSRPVLLQHLVGGASNNWISNPNNSLGFILADAGFDVWNGSRGN 135
 DB 61 YRIPYKKNSENIGKRPVAYLQHLIASATNWTNPNNSLAFILADAGFDVWNGSRGN 120
 QY 136 AWRKHKHTLSIDODEWAFSYDEMARFDPVAVNFILQKTQOEKIYVVGYSQGTMGFTA 195
 DB 121 TWSRKNVYSPDSVEFWAFSDEMAKYDLPATIDFVQKTQOEKIYVVGYSQGTMGFTA 180
 QY 196 FSTMPLEAKIKMYFALAPATATVKHAKSPGKTKELLDPDMIKGLFGKKEFL---YOTREFL 252
 DB 181 FSTNPALAKIKRYFALAPVATVKYTESPKKISLPKLLKVFIGNKMFPHNYLDQFL 240
 QY 253 RQLVYLCGQVILQICSNIMLLGGFNTNMNMSRASVYAAHTLAGTSVQNILHWSQAV 312
 DB 241 GTEV---CSRELLDLCLSNALFIFCGFDKKNLVSRFDVYLGHPNAGTSTQDPLFHAQLA 297
 QY 313 NSCELRAFDSWSEKLNKCNQPTPVRYRDMTVPYVAVNMGHDIADPQDVAMLLPKL 357
 DB 298 KSGKLAQYNGWSPLONLHYNQKTPPYDVVSAMTVPIAVWNGGHDIADPQDVAMLLPKL 357
 QY 373 TNLIIYHKNIPEWAHVDIFWGLDAPHRMYNEIHLMOOE 410
 DB 358 PNLLYHKKEILPYNHLDFIWAADPAQEVYNEIVTMMAED 395

RESULT 11

Q9D7C5 PRELIMINARY; PRT; 395 AA.
 AC Q9D7C5;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE 2310051B21RIK protein.
 GN 2310051B21RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TONGUE;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Offelli D., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sakai K., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK009359; BAB26240.1; -;
 DR MGD: MGI:1914967; 2310051B21RIK.
 DR InterPro: IPR000073; Abhydrolase.
 DR InterPro: IPR000734; Lipase.
 DR Pfam: PF00561; abhydrolase; 1.
 DR PROSITE: PS00120; LIPASE_SER; 1.
 SQ SEQUENCE 395 AA; 47478 MW; 3ED43992458DE058 CRC64;

Query Match 50.2%; Score 1127.5; DB 11; Length 395;
 Best Local Similarity 53.3%; Pred. No. 9.1e-91;

Matches 212; Conservative 69; Mismatches 110; Indels 7; Gaps 3;
 QY 17 MWLLIIVAYMFO-RNYSVHMPTKAYDPEAFMNISETIOHOGYCEPEYEVATEDGYLSV 75
 DB 1 MWLLIIVTSVLSAFGGAGHLEGLKPRNPEANMNVSMQITWGPSEYEVATEDGYLGV 60
 QY 76 NRIPRGLVOPKTKGSRPVLLQHLVGGASNNWISNPNNSLGFILADAGFDVWNGSRGN 135
 DB 61 YRIPYKKNSENIGKRPVAYLQHLIASATNWTNPNNSLAFILADAGFDVWNGSRGN 120
 QY 136 AWRKHKHTLSIDODEWAFSYDEMARFDPVAVNFILQKTQOEKIYVVGYSQGTMGFTA 195
 DB 121 TWSRKNVYSPDSVEFWAFSDEMAKYDLPATIDFVQKTQOEKIYVVGYSQGTMGFTA 180
 QY 196 FSTMPLEAKIKMYFALAPATATVKHAKSPGKTKELLDPDMIKGLFGKKEFL---YOTREFL 252
 DB 181 FSTNPALAKIKRYFALAPVATVKYTESPKKISLPKLLKVFIGNKMFPHNYLDQFL 240
 QY 253 RQLVYLCGQVILQICSNIMLLGGFNTNMNMSRASVYAAHTLAGTSVQNILHWSQAV 312
 DB 241 GTEV---CSRELLDLCLSNALFIFCGFDKKNLVSRFDVYLGHPNAGTSTQDPLFHAQLA 297
 QY 313 NSCELRAFDSWSEKLNKCNQPTPVRYRDMTVPYVAVNMGHDIADPQDVAMLLPKL 357
 DB 298 KSGKLAQYNGWSPLONLHYNQKTPPYDVVSAMTVPIAVWNGGHDIADPQDVAMLLPKL 357
 QY 373 TNLIIYHKNIPEWAHVDIFWGLDAPHRMYNEIHLMOOE 410
 DB 358 PNLLYHKKEILPYNHLDFIWAADPAQEVYNEIVTMMAED 395

RESULT 12

Q9D766 PRELIMINARY; PRT; 395 AA.
 AC Q9D766;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE 2310051E21RIK protein.
 GN 2310051E21RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TONGUE;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Offelli D., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sakai K., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK009544; BAB26350.1; -;
 DR MGD: MGI:1914967; 2310051B21RIK.
 DR InterPro: IPR000073; Abhydrolase.
 DR InterPro: IPR000734; Lipase.
 DR InterPro: IPR000379; Ser_estrs_site.

```
DR Pfam; PF00561; abhydrolase; 1.  
DR PROSITE; PS00120; LIPASE_SER; 1.  
SQ SEQUENCE 395 AA: 44648 MW; CC69875653AA7A7A CRC64;  
  
Query Match 50.1%; Score 1126.5; DB 11; Length 395;  
Best Local Similarity 53.3%; Pred. No. 1.1e-90;  
Matches 212; Conservative 70; Mismatches 109; Indels 7; Gaps 3;  
  
QY 17 MWLLILVAYMFO-RNVNSVHMTKAVDPEAFWNISIIIOHOGYPCPEEYEVATEDGYLSV 75  
DB 1 MWLLLVTSVLSAFGGAHGLFGKLGPKNPANNVNSOMITYWGPSEYEVATEDGYILGV 60  
  
QY 76 NRIPRGLVOPKTKGRPVVLLQHLGVGGASNNISLNPNNSLGFLADAGDFVMMGNSRGN 135  
DB 61 YRIPYKKNSEINIGKRPVAYLQHLIASATNTITLNPNNSLAFILADAGYDVLGNSRGN 120  
  
QY 136 ANSRKHTLSIDODEFAESYDEMARFDLPVINFILQKTGQEKIYYVYSGTGTGFTA 195  
DB 121 TWSRKNVYSPDSVEFAESFDEMAKYDLPATIDFVQKTGQEKIHYVGHSGTGTGFTA 180  
  
QY 196 FSTMPELAQKIRKMFALAPVATVVKHAKSPGTFFLLPDMNKGFLGKKEFL---YOTREL 252  
DB 181 FSTNPALAKKIRKMFALAPVATVVKYTESPFKISLIPKELLVIFGNKMFMPYNYLDQEL 240  
  
QY 253 ROLVYILCOVILDOICSNIMLLGFGNTNNMNSRASVYAAHTLAGTSVQNTLHWSQAV 312  
DB 241 GTEV---CSRELLDLCSNALFIFCGDFKKNLNSRFDVYLGNHPAGTSTQDLFHAQLA 297  
  
QY 313 NSGELRAFQWSETKLEKNQPTPVRYVRDMTPTAMWTGQDMLSPEDVKMLLSEV 372  
DB 298 KSGKLOAYNWGSPQLNHLHYNOKTPPYDVVSAMTVPVIAVWNGHDLADPQDVAMLLPKL 357  
  
QY 373 TNLIIYHKNIPEWAHVDFIWLGDAPHRYNEIHLMOQE 410  
DB 358 PNLLYHKELIPYNHLDFIWMDAPQEVYNEIVTMAED 395  
  
RESULT 13  
Q9D6L1  
ID Q9D6L1 PRELIMINARY; PRT; 395 AA.  
AC Q9D6L1;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Adult male tongue cDNA, RIKEN full-length enriched library,  
DE clone:2310079020, full insert sequence.  
GN 2310051B21RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=TONGUE;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Glasi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohtsuki S.,  
RA Hayashizaki Y.;
```

```
RT *Functional annotation of a full-length mouse cDNA collection.*;  
RL Nature 409:685-690(2001).  
DR EMBL; AK010231; BAB26784.1; -;  
DR MGD; MGI:1914967; 2310051B21RIK.  
DR InterPro; IPR000073; Abhydrolase.  
DR InterPro; IPR000734; Lipase.  
DR InterPro; IPR000379; Ser_estr_site.  
DR Pfam; PF00561; abhydrolase; 1.  
DR PROSITE; PS00120; LIPASE_SER; 1.  
SQ SEQUENCE 395 AA: 44659 MW; 13E0BB95FC5A40E0 CRC64;  
  
Query Match 50.1%; Score 1126.5; DB 11; Length 395;  
Best Local Similarity 53.5%; Pred. No. 1.1e-90;  
Matches 213; Conservative 69; Mismatches 109; Indels 7; Gaps 3;  
  
QY 17 MWLLILVAYMFO-RNVNSVHMTKAVDPEAFWNISIIIOHOGYPCPEEYEVATEDGYLSV 75  
DB 1 MWLLLVTSVLSAFGGAHGLFGKLGPKNPANNVNSOMITYWGPSEYEVATEDGYILGV 60  
  
QY 76 NRIPRGLVOPKTKGRPVVLLQHLGVGGASNNISLNPNNSLGFLADAGDFVMMGNSRGN 135  
DB 61 YRIPYKKNSEINIGKRPVAYLQHLIASATNTITLNPNNSLAFILADAGYDVLGNSRGN 120  
  
QY 136 ANSRKHTLSIDODEFAESYDEMARFDLPVINFILQKTGQEKIYYVYSGTGTGFTA 195  
DB 121 TWSRKNVYSPDSVEFAESFDEMAKYDLPATIDFVQKTGQEKIHYVGHSGTGTGFTA 180  
  
QY 196 FSTMPELAQKIRKMFALAPVATVVKHAKSPGTFFLLPDMNKGFLGKKEFL---YOTREL 252  
DB 181 FSTNPALAKKIRKMFALAPVATVVKYTESPFKISLIPKELLVIFGNKMFMPYNYLDQEL 240  
  
QY 253 ROLVYILCOVILDOICSNIMLLGFGNTNNMNSRASVYAAHTLAGTSVQNTLHWSQAV 312  
DB 241 GTEV---CSRELLDLCSNALFIFCGDFKKNLNSRFDVYLGNHPAGTSTQDLFHAQLA 297  
  
QY 313 NSGELRAFQWSETKLEKNQPTPVRYVRDMTPTAMWTGQDMLSPEDVKMLLSEV 372  
DB 298 KSGKLOAYNWGSPQLNHLHYNOKTPPYDVVSAMTVPVIAVWNGHDLADPQDVAMLLPKL 357  
  
QY 373 TNLIIYHKNIPEWAHVDFIWLGDAPHRYNEIHLMOQE 410  
DB 358 PNLLYHKELIPYNHLDFIWMDAPQEVYNEIVTMAED 395  
  
RESULT 14  
Q9D6Q3  
ID Q9D6Q3 PRELIMINARY; PRT; 395 AA.  
AC Q9D6Q3;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Adult male tongue cDNA, RIKEN full-length enriched library,  
DE clone:2310068C02, full insert sequence.  
GN 2310051B21RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=TONGUE;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Glasi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohtsuki S.,  
RA Hayashizaki Y.;
```

```
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010103; BAB26703.1;
DR MGI; MGI:1914967; 2310051B21Rik.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44638 MW; 4E52613BC98AB3F CRC64;

Query Match 50.18; Score 1125.5; DB 11; Length 395;
Best Local Similarity 53.3%; Pred. No. 1.4e-90;
Matches 212; Conservative 69; Mismatches 110; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFO-RNVSVMHPTKAVDPEAFMNISEIIIOHGYPCPEEVEVATEDGYILSV 75
DB 1 MWLLIVTSVLSAFSGAGHGLFGKLPKNEANNVSMITYWGYPSEYEVVTEGYILGV 60
QY 76 NRIPRGLVQPKTKTSRPPVYLQHLGVGGASNMISLNPNSLGLFILADAGFDVWNGSRGN 135
DB 61 YRIPYGRKNSNICKRPVAYLQHLGLIASATNWTNLPNSLAFILADAGYDVLGNSRGN 120
QY 136 AWSRKHHTLSIDQDEFWAFSDENARFDPVAVINFILOKTQOEKIYVYSGTGMFTG 195
DB 121 TWSRKNVYSPDSVEFWAFSDENAKYDLPATIDFIVQKTQOEKIHYVGHSGQTIGFIA 180
QY 196 FSTMPDLAOKTKMYFALAPIATVHKAKSPGTFKLLPDMIKLFGKKEFL---YQTRFL 252
DB 181 FSTNPALAKKIKRFYALAPVATVYTESPFKISLIPFKLVIFGNKMFPHNYLDQFL 240
QY 253 RQVLYLGGQVILQDQICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAV 312
DB 241 GTEV---CSRELLDLCNALFIFCGFDKKNLVSRFDVYXGHNPACTSTQDLFHWQA 297
QY 313 NSGELRAFDMGSETKNLEKCNQPTPVRYVRDMVTVPYVAMTGGQDWLSNPEDVKMLSEV 372
DB 298 KSGKLQAYNWGSPQLQNLHYNQKTPPYVDVSANTVPIAVWNGGHDILADPDQVAMLLPK 357
QY 373 TNLVYHKNIPEWAVHDFIWLGDAPHMYNEIHLWQOE 410
DB 358 PNLVYHKELPNHLYFIWAMDAPQEVYNEIVTMMAED 395

RESULT 15
Q9D760 PRELIMINARY; PRT; 395 AA.
AC Q9D760
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE 2310051B21Rik protein.
GN 2310051B21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Arizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
```

```
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
RA Schirai L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK009560; BAB26359.1;
DR MGI; MGI:1914967; 2310051B21Rik.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44685 MW; PEF96B65EA670BEE CRC64;

Query Match 50.08; Score 1124.5; DB 11; Length 395;
Best Local Similarity 53.3%; Pred. No. 1.7e-90;
Matches 212; Conservative 70; Mismatches 109; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFO-RNVSVMHPTKAVDPEAFMNISEIIIOHGYPCPEEVEVATEDGYILSV 75
DB 1 MWLLIVTSVLSAFSGAGHGLFGKLPKNEANNVSMITYWGYPSEYEVVTEGYILGV 60
QY 76 NRIPRGLVQPKTKTSRPPVYLQHLGVGGASNMISLNPNSLGLFILADAGFDVWNGSRGN 135
DB 61 YRIPYGRKNSNICKRPVAYLQHLGLIASATNWTNLPNSLAFILADAGYDVLGNSRGN 120
QY 136 AWSRKHHTLSIDQDEFWAFSDENARFDPVAVINFILOKTQOEKIYVYSGTGMFTG 195
DB 121 TWSRKNVYSPDSVEFWAFSDENAKYDLPATIDFIVQKTQOEKIHYVGHSGQTIGFIA 180
QY 196 FSTMPDLAOKTKMYFALAPIATVHKAKSPGTFKLLPDMIKLFGKKEFL---YQTRFL 252
DB 181 FSTNPALAKKIKRFYALAPVATVYTESPFKISLIPFKLVIFGNKMFPHNYLDQFL 240
QY 253 RQVLYLGGQVILQDQICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAV 312
DB 241 GTEV---CSRELLDLCNALFIFCGFDKKNLVSRFDVYXGHNPACTSTQDLFHWQA 297
QY 313 NSGELRAFDMGSETKNLEKCNQPTPVRYVRDMVTVPYVAMTGGQDWLSNPEDVKMLSEV 372
DB 298 KSGKLQAYNWGSPQLQNLHYNQKTPPYVDVSANTVPIAVWNGGHDILADPDQVAMLLPK 357
QY 373 TNLVYHKNIPEWAVHDFIWLGDAPHMYNEIHLWQOE 410
DB 358 PNLVYHKELPNHLYFIWAMDAPQEVYNEIVTMMAED 395

Search completed: February 3, 2003, 13:49:58
Job time : 37 secs
```